Enabling Connections in the Product Lifecycle using the Digital Thread

Thomas D. Hedberg, Jr.

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Industrial and Systems Engineering

Jaime A. Camelio, Chair
Konstantinos P. Triantis
Alejandro Salado Diez
Brian A. Weiss

October 12, 2018
Falls Church, Virginia

Keywords: Product Lifecycle Management, Digital Thread, Smart Manufacturing

Enabling Connections in the Product Lifecycle using the Digital Thread

Thomas D. Hedberg, Jr.

(ABSTRACT)

Product lifecycles are complex heterogeneous systems. Applying control methods to lifecycles requires significant human capital. Additionally, measuring lifecycles relies primarily on domain expertise and estimates. Presented in this dissertation is a way to semantically represent a product lifecycle as a cyber-physical system for enabling the application of control methods to the lifecycle. Control requires a model and no models exist currently that integrate each phase of lifecycles. The contribution is an integration framework that brings all phases and systems of a lifecycle together. First presented is a conceptual framework and technology innovation. Next, linking product lifecycle data dynamical is described and then how that linked data could be certified and traced for trustworthiness. After that, discussion is focused how the trusted linked data could be combined with machine learning to drive applications throughout the product lifecycle. Last, a case study is provided that integrates the framework and technology. Integrating all of this would enable efficient and effective measurements of the lifecycle to support prognostic and diagnostic control of that lifecycle and related decisions.

Enabling Connections in the Product Lifecycle using the Digital Thread

Thomas D. Hedberg, Jr.

(GENERAL AUDIENCE ABSTRACT)

The manufacturing sector is on a precipice to disruptive change that will significantly alter the way industrial organizations think, communicate, and interact. Industry has been chasing the dream of integrating and linking data across the product lifecycle and enterprises for decades. However, inexpensive and easy to implement technologies to integrate the people, processes, and things across various enterprises are still not available to the entire value stream. Industry needs technologies that use cyber-physical infrastructures effectively and efficiently to collect and analyze data and information across an enterprise instead of a single domain of expertise. Meeting key technical needs would save over $100 billion annually in emerging advanced manufacturing sectors in the US. By enabling a systems-thinking approach, significant economic opportunities can be achieved through an industrial shift from paper-based processes to a digitally enabled model-based enterprise via the digital thread. The novel contribution of this dissertation is a verified and validated integration framework, using trusted linked-data, that brings all phases and systems of the product lifecycle together. A technology agnostic approach was pursued for dynamically generating links. A demonstration is presented as a reference implementation using currently available technology. Requirements, models, and policies were explored for enabling product-data trustworthiness. All methods were developed around open, consensus-based standards to increase the likelihood of scalability. The expected outcome of this work is efficient and effective measurements of the lifecycle to support data-driven methods, specifically related to knowledge building, decision support, requirements management, and control of the entire product lifecycle.

Dedication

To all my Product Lifecycle Management (PLM) brothers and sisters who put in a good fight trying to get some work done each and every day.

Acknowledgments

In speaking with others who have pursued a Ph.D., there is always an underestimation in the amount of effort required to finish. I thankfully had a strong support group. I would like to honor a few colleagues and individuals who helped realized this work. First, I wish to thank Dr. William Regli for providing the motivation and guidance to pursue the Lifecycle Information Framework and Technology (LIFT) concept. Dr. Regli pointed me to other domains with similar ideas and enabled me to discover key lessons learned from those endeavors. Next, many thanks must be extended to my NIST office mate, partner is causing disruption, and friend, Dr. Moneer Helu, for the many late evening ”fire-side” chats in the office. The discussions with Dr. Helu helped mature the LIFT concept into a viable product. I must also thank Ms. Allison Barnard Feeney for becoming my “forced” copy editor through the NIST Editorial Review Board process. Ms. Barnard Feeney taught me that “deleting every other word” does make the manuscript better. ;-) Thank you to Dr. Sylvere Krima for all the help with coding the various applications developed during this research adventure, especially the Digital Manufacturing Certificates toolkit. Thank you to Mr. Curtis Brown and Mr. Adrian Miura for adopting and testing the toolkit. Mr. Brown’s and Mr. Miura’s use of the toolkit in a real industrial setting provided significant and high-quality feedback that made the work better. Thank you to Intercax for adopting the link-data methodology and commercializing portions of it. To my NIST, academic colleagues, and industry friends, thank you for your comments, reviews, feedbacks, and general interest in my work. Thank you to my dissertation committee for providing the guidance and mentoring to get this thing done. Lastly, thank you to my family for putting up with my occasional grumpiness and constant lack of free time. Becky, Ethan, and Lucas, you all provided the motivation and understanding that this struggle is temporary and our family will all be better off because of the sacrifice.

[0]This work was supported in part by the National Institute of Standards and Technology (NIST). Certain commercial systems may be identified in this paper. Such identification does not imply recommendation or endorsement by NIST. Nor does it imply that the products identified are necessarily the best available for the purpose.

Contents

List of Figures

List of Tables

List of Algorithms

List of Listings

Chapter 1

Introduction

1.1 Overview

The product lifecycle is a complex heterogeneous system. Applying control methods to the lifecycle requires significant human capital. Additionally, measuring the lifecycle relies primarily on domain expertise and estimates. In this research, I present a way to semantically represent the product lifecycle as a cyber-physical system supported by the digital thread. Digital Thread is an integrated information flow that connects all of the phases of the product lifecycle using an accepted authoritative data source (e.g., technical data package [1], three-dimensional (3D) computer-aided design (CAD) model) [2, 3, 4]. The contribution of this dissertation is a verified and validated integration framework, using trusted linked-data, that brings all phases and systems of the product lifecycle together. This would enable efficient and effective measurements of the lifecycle to support data-driven methods, specifically related to knowledge building, decision support, requirements management, and control.

The most advantageous time to make a change in a product's lifecycle is in the early stages because there is an inherent flexibility in those decisions being made [5] and the cost of changes is less than having to make changes later in the lifecycle. Ullman [6] said more knowledge is gained about a product or process as time and the lifecycle progress, but less freedom to make changes is available later in the lifecycle because of costs and other constraints. This is known as the "design paradox." Salado and Nilchiani [7], Salado et al. [8] showed that every decision made early in the lifecycle becomes a constraint on the remainder of the lifecycle, because each subsequent decision further reduces the compliant solution space.

Several concepts and methods exist to help make early decisions in the product lifecycle. However, these concepts, such as Total Design theory [9], require prerequisite knowledge of the various phases of the product lifecycle. In a sampling of 35 defense-acquisition programs [10], development-cost growth averaged 57 percent and procurement-cost growth averaged 75 percent. Decisions dominated the growth in both types of cost growth. These are the reasons that motivate pursuing a concept that could support prognostic and diagnostic control of decisions in the product lifecycle.

Collaborative Product Development [11], Concurrent Engineering [12], Designed for Manufacturing [13, 14], Design for Six Sigma [15], and Integrated Product and Process Development [16] are popular business strategies industry uses for managing new-development activities. Decision making is a common function in all of these strategies. Companies may combine these popular strategies with stage-gate processes to form their operating models. Further, industry desires to couple these methods with model-based systems engineering (MBSE), the "vee" diagram, and the larger-scoped model-based enterprise (MBE) concept to enable effective decision making during development and manufacturing processes [17].

However, organizations often apply these methods without ever re-asking if the development and manufacturing activities are still the right pursuits – that is, should the organization's overall goals change during and throughout the activities? This question and the desire to ensure the optimality, stability, effectiveness, and efficiency of technological-innovation process motivates this work.

Smart manufacturing cannot be successful without proper management and technological innovation in decision making. The Oxford English Dictionary [18] defines technology as "the application of such knowledge for practical purposes." Innovation [19] is defined as "the alteration of what is established by the introduction of new elements or forms." And, management [20] is defined as "organization, supervision, or direction." Using these definitions, technological innovation may be defined as the process for creating a new application of practical knowledge. Thus, the management of technological innovation is the organization, supervision, or direction of the process for creating a new application of practical knowledge.

While, several "smart" technologies have existed in manufacturing since the 1980s [21, 22], the integration of those technologies along with the convergence of information technology (IT) and operational technology (OT) has kicked off a period of an increased rate of innovation in manufacturing. In general, Tidd and Bessant [23] presented "key lessons learned about managing innovation." Tidd and Bessant [23] recommended that organizations be visible in promoting innovation across the whole business, build a project-based organization with a good portfolio management structure, utilize a stage-gate system, and institutionalize the use of tools. Remember, innovation requires the creation of something new. Therefore, creativity, development processes, and change management must be accounted for in decision making within the overall technological-innovation process.

Success criteria for the research is to enable a reference implementation of an integration framework for automating the diagnosing of decisions in support of building new knowledge. Figure 1.1 represents the decision-making process based on generated knowledge and experience. In the decision-making process, shown in Figure 1.1, the knowledge box represents a group of "answers" to previous, and even future, questions. The decision box represents the recognition of a question. Engineers use their knowledge to make decisions. Unfortunately, that is often where the process stops. Due to interactions and pressures in the work environment, a secondary question of, "how good was the decision?" is not often addressed – even if it is recognized as an important question.

While the quality of a decision does not need to be analyzed fully, at some point during

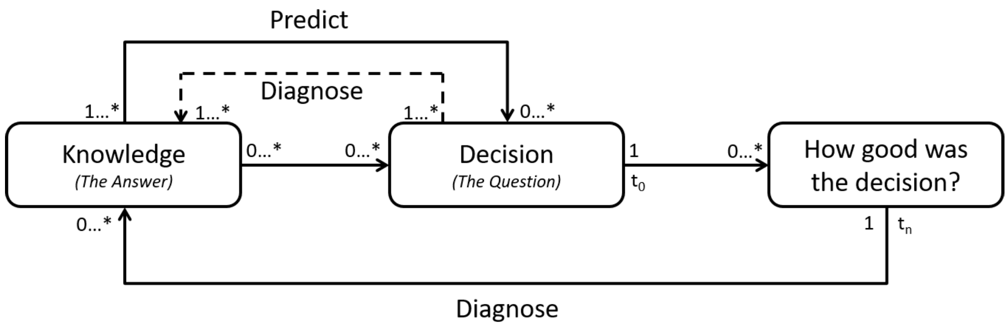


Figure 1.1: The knowledge-decision cycle for knowledge-driven decision making

the product lifecycle processes, decisions should be reviewed. If t_0 is a point in time when a decision is made, then at t_n, where n is simply some future point in time, the quality of the decision could be diagnosed and classified as good or bad. If the affects of the decision have not played out completely, the diagnosis of the decision will be made under some level of uncertainty. Quantifying the uncertainty of the decision diagnosis is out of scope for the this research, but one must remain cognizant of the fact that there is a level of uncertainty in the diagnosis process.

Using diagnosis and prognosis methods to support knowledge-driven decision making would be a novel contribution to the field. The concept in Figure 1.1 differs from traditional "big data" methods or data-driven approaches because the common approach today is to collect data before really considering the decisions. This is analogous to answering a question before knowing what is the question. The common approach hinders the ability to effectively use data-driven methods in real time without a considerable amount of pre-processing and training. However, the integration framework, combined with the concept depicted in Figure 1.1, could support better linking of data to assist in diagnosing decisions. This would enable more informed decision making and considerations of the decisions that need to be made.

The contribution of this research is an integration framework to address three industry needs: (1) linking data from and across the lifecycle phases, (2) ensuring authentication, authorization, and traceability of product data, and (3) integrating feed-back and feed-forward process control of product lifecycle processes. New methods and technology were developed for verifying and validating the integration framework in support of the three objectives. The integration framework was verified and validated using a real-world case study focused on data-driven decisions using data from the product lifecycle.

1.2 Research Problem

This research developed and evaluated the integration framework presented in Chapter 4. This framework relies on open-data standards and generally accepted practices from the information science and product lifecycle management (PLM) domains. A successful reference implementation of this framework demonstrates effective and efficient curation, discovery, and reuse of distributed data without the need of expensive large-enterprise data-management platforms. The following inquiry-based questions helped form the research questions for the dissertation work:

- What technologies (e.g., commercial tools, standards, open-source) exist for linking data from and across the lifecycle phases?
- How can information models from different domains be dynamically (e.g., on-demand) linked?
- How can different viewpoints of the same data and/or information be reconciled in near realtime?
- How can the product lifecycle be connected via vertical and horizontal integrations (e.g., vertically up and down the ISA-95 pyramid and horizontally across multiple domains)?
- What are the standard information interfaces between each phase of the product lifecycle?
- Can a standard method for information-model and viewpoint interoperability be defined?
- What technologies exist for enabling authentication, authorization, and traceability of product data?
- Does product data quality (PDQ) affect authentication, authorization, and traceability?
- What is a method for supporting asynchronous and synchronous data authentication, authorization, and traceability?
- What technologies exist for integrating feed-back and feed-forward process control of product lifecycle processes?
- What methods exist for applying machine learning / artificial intelligence to generate knowledge bases and support manufacturing-related decisions?
- Are there gaps in control and decision theory that limit the ability to utilize the LIFT

concept? If so, what are they?

Answers to the inquiry generated much of the background knowledge presented in Chapter 3. The process of the inquiry and the resultant background knowledge led to the discovery of several research questions. In general, the overarching question that must be answered is, *how can the product lifecycle be represented semantically as a cyber-physical system supported by the digital thread to enable monitoring and measuring the lifecycle for applying control methods at the enterprise level of the lifecycle?* However, to achieve the goal of applying the control methods to the entire product lifecyle and thus answering the overarching research question, research must be completed in well-scoped phases. Therefore, the following research questions are proposed for answering as the first steps towards achieving the stated goals:

1. How can Graph Theory be used to dynamically generate inter-domain links of product data between the phases of the product lifecycle?

2. How can linked product data be managed to enable authentication and authorization as the data moves between domains?

3. What kind of management and/or governance policy could be established to ensure trustworthiness of linked product data?

Thus, it is hypothesized that developing and evaluating the integration framework proposed in this research would answer the overarching research question. Control requires a model and no models exist currently that integrate each phase of the product lifecycle. The contribution of an integration framework that brings all phases and systems of a lifecycle together would enable developing the required models and measurement system for monitoring the lifecycle. The three proposed research questions for this dissertation work is the beginning down a path towards a large scale effort of fundamentally changing how product lifecycle management is approached. The work plan used in answering the proposed research questions is presented in Section 1.4.

1.3 Aims and Objectives

The aim of this work was to develop an integration framework that brings all phases and systems of a product lifecycle together to enable efficient and effective measurements of the lifecycle in support of data-driven methods, specifically related to knowledge building, decision support, requirements management, and control.

To achieve the aim, the following objectives are defined:

1. Link universally heterogeneous information systems and data sets across the various domains of the product lifecycle dynamically without requiring one-to-one data mapping

2. Utilize cyber-security and cryptography technologies in novels ways to support data and system trustworthiness with authentication, authorization, and traceability of product data

3. Enable data-driven-application use cases, using trusted linked data, support domain-specific knowledge base development, requirements management, decision support, and feed-back and feed-forward process control

The above objectives were set toward the development of an integration framework to guide the development of dynamic integrations of data and process management systems across the product lifecycle. The product lifecycle is studied as a cyber-physical system. Additionally, the relationship between the digital artifacts and physical processes of the lifecycle is investigated.

1.4 Plan of Work

To complete the dissertation, the research pursued a conceptual framework and technology innovation. First, the research investigated if and how Graph Theory supports linking product lifecycle data dynamically. Next, the research studied how that linked data would be certified and traced for trustworthiness within the proposed policies will be investigated. Lastly, a management and governance policy was developed using the certified, linked product data for establishing trustworthy interactions within and between phases of the lifecycle. Integrating all of this would start down the path of enabling efficient and effective measurements of the lifecycle to support prognostic and diagnostic control of that lifecycle and related decisions.

Each phase of the research started with a literature review to determine the current state of the art. The research developed systems and information models at different levels of abstraction. Those models were tested through simulation and prototype manufacturing environments to verify and validate them. The research leveraged the National Institute of Standards and Technology (NIST) Smart Manufacturing Systems Test Bed (SMS Test Bed) [1] for testing physical processes. The overall framework was verified and validated through a case study described in Section 5.3.

A technology agnostic approach was pursued to answer the first research question on dynamically generating links. Then, the approach was demonstrated as a reference implementation using a currently available technology. Then, research commenced through requirements gathering for trustworthiness related to interacting with product data. Regulated industries (e.g., Federal Aviation Administration (FAA), Food and Drug Administration (FDA)) will be reviewed to determine any statutory requirements. Next, a model was developed to propose a policy for enabling product-data trustworthiness. The policy model will be integrated with

[1] https://smstestbed.nist.gov

the answer from the first research question to ensure compliance. Lastly, a data-traceability method was developed. The research strived to develop the traceability method around open, consensus-based standards to increase the likelihood of scalability. The traceability method was demonstrated using ISO 10303-21 Edition 3 (the Standard for the Exchange of Product Model Data (STEP) EXPRESS implementation form) and Quality Information Framework (QIF) standards, which are two popular standards used by industry. Answering the three research questions starts the domain down the path of applying control methods at the enterprise level of the lifecycle. Thus, the answers to the research questions provides novel theoretical contributions to the PLM domain.

1.5 Original Contribution

This section will outline the expected outcomes and potential impacts of the research. The outcomes are described in the context of contributions to the manufacturing domain. The impacts are described in the context of the novel content of the research.

1.5.1 Contribution to the manufacturing domain

The main contribution the research is making to the manufacturing domain is a verified and validated framework and technology infrastructure for the digital thread. The literature supports the need of a formalized digital-thread framework. Much of the literature is focused on the digital twin – the generation of surrogate models in the cyber-space to replicate the physical product. However, the digital thread, which enables the digital twin, has received minimal attention in the literature. Industry and solution providers are searching for a common technology approach to enable the digital thread. The proposed research would close this gap.

Further, as evident in the U.S. Department of Defense (DoD) Military Engineering Data Asset Locator System (MEDALS) [24], the amount of human capital required to keep the data repositories up-to-date is significant. The lifecycle can no longer afford to deploy the needed amount of human capital to maintain near-dynamic knowledge bases and still deliver products competitively to market. Industry needs a way to discover data relationships and link the data across the product lifecycle. Therefore, trusted linked data is another contribution the proposed research brings to the manufacturing domain.

Trusted linked data would enable near-real-time dynamic updating of domain-specific knowledge bases in the product lifecycle by using machine learning and artificial intelligence methods. With the proposed Lifecycle Information Framework and Technology (LIFT) concept, the only human capital required to enable automated updating would be a data administrator needs to register a domain-specific knowledge base with a “registry” to ensure data is discoverable. Once that registration is complete, the knowledge base would receive near-real-time dynamic updates as users complete their day-to-day activities.

The LIFT concept would enable input of all the disparate pieces of data, from information silos in the lifecycle, into a self-learning and self-aware system supported by data-driven applications. The system could utilize self-learning algorithms for looking at semantic, syntactic, linguistic, and statistic data that comes from all the different data and information repositories of design, analysis, manufacturing, quality, and customer and product support. The stream of data and information into these self-learning algorithms could enable the system to learn dynamically from streams of data based on product experience. The learning supports near-real-time dynamic updating of design knowledge bases in an effective manner. The effective and traceable information flow, self-learning methods, dynamic updating of the design knowledge base, and real-time decision support form the framework of the LIFT concept.

1.5.2 Novel content of the research

Aside from the lifecycle information framework, several novel concepts are included in this research proposal. While many of the technologies (e.g., PLM, knowledge management, decision support) discussed in this proposal or required to complete the research are not novel, the integration of those technologies is novel. There is no evidence in the literature to support that prior work tried or was successful in completely integrating technologies across the product lifecycle.

In addition, the systematic literature review (SLR) discussed in Chapter 2 provides evidence that the digital thread is an emerging concept. Industry has long sought to link together data from different sources in the lifecycle. However, incompatible information models and systems remain a barrier to achieving industry's need. The proposal to integrate semantic web and linked-data techniques into manufacturing is a novel contribution that could support closing the gap and enable industry to curate, discover, and retrieve data more effectively and efficiently across the product lifecycle.

The proposed approach to data certification and traceability to support trustworthiness is another novel approach that builds on the proposed linked-data work. While the X.509 [25] standard has been published since 1988 and revised several times since, the use of X.509 for authorization, authentication, and traceability of data in manufacturing is a novel and simple approach. Historically, X.509 has been used in manufacturing for cyber-security and encrypting data. But, the approach the proposed research would extend the use of X.509 to enable trustworthy storage and exchange of data in manufacturing. The extension would support industry in meeting traceability requirements from regulatory agencies such as FAA and FDA.

Lastly, the backend schematic (see Figure 4.2) of the LIFT concept is a novel approach to support the digital thread infrastructure. The application of micro-services architecture to

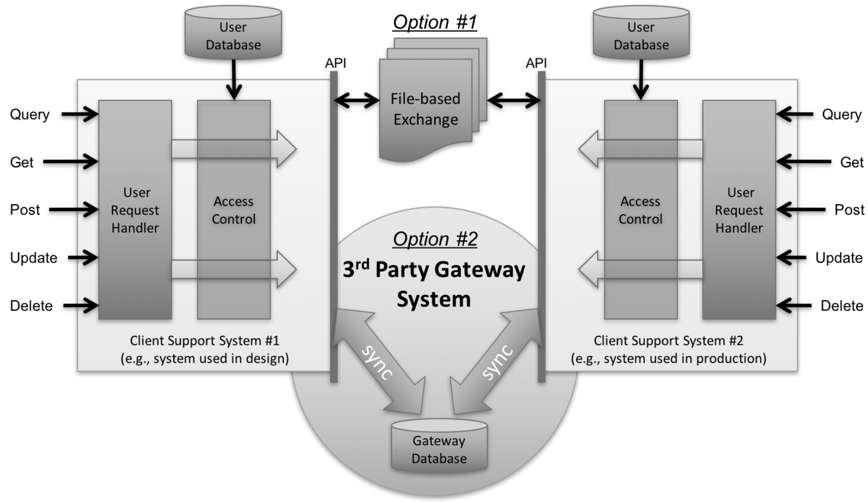


Figure 1.2: Current state of the art for exchanging data across the lifecycle using one-to-one copies of data between systems.

the enterprise in support of universal enterprise-wide query and object management is a unique contribution to manufacturing that has not been explored fully. Today industry uses two prevalent methods for exchanging data between systems: (1) file-based exchange and (2) gateway systems. Figure 1.2 provides a pictorial overview of each exchange method used in industry today. In file-based exchanges, data is copied between systems using proprietary "packaging" tools or a standard such as ISO 10303-239 [26] (also known as PLCS). In gateway exchanges, data is copied between systems to a centralized repository and then each system connected to the central repository is synced.

In both exchange methods, data is copied one-to-one between all systems. Data can quickly become out of sync and discrepancies creep into the lifecycle, which reduces the quality of data available to users. The proposed research would develop a decentralized method for exchanging data – where the data always remains where the data was authored and connected systems would be able to retrieve a view of the data for exchange and collaboration purposes. This would ensure that the data user is always seeing the most up-to-date version of the data.

1.6 Layout of the Dissertation

The following provides an overview for the layout of this dissertation. A breif explanation of each chapter is included in the list below.

1. Introduction

2. Literature Review: provides a systematic literature review of digital thread and digital twin to identify areas of needed research

3. Background: provides additional background knowledge on the domains addressed by the research

4. Conceptual Framework and Technology Innovation: provides a high-level overview of the developed "Lifecycle Information Framework and Technology" delivered by the research

5. Linked Data: describes using graphs for linking data across the product lifecycle to enable digital threads

6. Data certification and traceability: explains the digital manufacturing certificates toolkit and architecture for enabling digital certification and traceability of product-related data

7. Lifecycle of trust: details trust methodology supported by a structure and governance policy to support trustworthiness through the product lifecycle

8. Conclusion: provides some concluding analysis and proposes future work

1.7 Publications

Part of the contents of this dissertation has been submitted and/or accepted for publication. The publications as of the approval of this dissertation are listed below:

- Hedberg Jr, T., Krima, S., & Camelio, J. A. (2016). Embedding X.509 Digital Certificates in Three-Dimensional Models for Authentication, Authorization, and Traceability of Product Data. *Journal of Computing and Information Science in Engineering*, 17(1), 011008-011008-011011. doi:10.1115/1.4034131

- Hedberg Jr, T., Barnard Feeney, A., Helu, M., & Camelio, J. A. (2017). Towards a Lifecycle Information Framework and Technology in Manufacturing. *Journal of Computing and Information Science in Engineering*, 17(2), 021010-021010-021013. doi:10.1115/1.4034132

- Hedberg Jr, T., Barnard Feeney, A., & Camelio, J. A. (2018). Towards a Diagnostic and Prognostic Method for Knowledge-Driven Decision Making in Smart Manufacturing Technologies. In A. M. Madni, B. Boehm, R. G. Ghanem, D. Erwin, & M. J. Wheaton (Eds.), *Disciplinary Convergence in Systems Engineering Research*: Springer International Publishing.

- Hedberg Jr, T., Krima, S., & Camelio, J. A. (In Review). Method for Enabling a Root of Trust in Support of Product-Data Certification and Traceability. *Journal of Computing and Information Science in Engineering*

- Hedberg Jr, T., Bajaj, M., & Camelio, J. A. (In Review). Using Graphs to Link Data Across the Product Lifecycle for Enabling Smart Manufacturing Digital Threads. *Journal of Computing and Information Science in Engineering*

Chapter Bibliography

[1] US Department of Defense. Standard practice: Technical data packages, MIL-STD-31000 rev. A., 11/1/2009 2013.

[2] Edward M. Kraft. The US Air Force digital thread/digital twin life cycle integration and use of computational and experimental knowledge. In *54th AIAA Aerospace Sciences Meeting, 2016, January 4, 2016 - January 8, 2016*. American Institute of Aeronautics and Astronautics, 2016.

[3] Thomas D. Hedberg Jr, Joshua Lubell, Lyle Fischer, Larry Maggiano, and Allison Barnard Feeney. Testing the digital thread in support of model-based manufacturing and inspection. *Journal of Computing and Information Science in Engineering*, 16(2): 1–10, 2016. ISSN 1530-9827. doi: 10.1115/1.4032697.

[4] Rivai Wardhani and Xun Xu. Model-based manufacturing based on STEP AP242. In *2016 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pages 1–5, 2016. doi: 10.1109/MESA.2016.7587187.

[5] Christian Johansson, Ben Hicks, Andreas C. Larsson, and Marco Bertoni. Knowledge maturity as a means to support decision making during product-service systems development projects in the aerospace sector. *Project Management Journal*, 42(2):32–50, 2011. ISSN 87569728. doi: 10.1002/pmj.20218.

[6] David G. Ullman. *The mechanical design process*. McGraw-Hill series in mechanical engineering. McGraw-Hill Higher Education, Boston, 4th edition, 2010.

[7] Alejandro Salado and Roshanak Nilchiani. A research on measuring and reducing problem complexity to increase system affordability: From theory to practice. *Procedia Computer Science*, 44:21–30, 2015. ISSN 1877-0509. doi: 10.1016/j.procs.2015.03.037.

[8] Alejandro Salado, Roshanak Nilchiani, and Dinesh Verma. A contribution to the scientific foundations of systems engineering: Solution spaces and requirements. *Journal of Systems Science and Systems Engineering*, 2016. ISSN 1861-9576. doi: 10.1007/s11518-016-5315-3.

[9] Stuart Pugh. *Total design : integrated methods for successful product engineering*. Addison-Wesley Pub. Co., Wokingham, England ; Reading, Mass., 1991. ISBN 0201416395.

[10] Joseph G. Bolten, Robert S. Leonard, Mark V. Arena, Obaid Younossi, and Jerry M. Sollinger. Sources of weapon system cost growth: Analysis of 35 major defense acquisition programs. Report, RAND Corporation, 2008. URL http://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND_MG670.pdf.

[11] Min Li, Shuming Gao, and Charlie C. Wang. Real-time collaborative design with heterogeneous CAD systems based on neutral modeling commands. *Journal of Computing and Information Science in Engineering*, 7(2):113–125, 2006. ISSN 1530-9827. doi: 10.1115/1.2720880.

[12] Andrew Kusiak. *Concurrent engineering : automation, tools, and techniques*. Wiley, New York, 1993. ISBN 0471554928 (cloth acid-free paper).

[13] Carlos Eduardo Sanches da Silva, Eduardo Gomes Salgado, Carlos Henrique Pereira Mello, Eduardo da Silva Oliveira, and Fabiano Leal. Integration of computer simulation in design for manufacturing and assembly. *International Journal of Production Research*, 52(10):2851–2866, 2014. ISSN 0020-7543. doi: 10.1080/00207543.2013.853887.

[14] Ibrahim H. Garbie. DFSME: design for sustainable manufacturing enterprises (an economic viewpoint). *International Journal of Production Research*, 51(2):479–503, 2013. ISSN 0020-7543. doi: 10.1080/00207543.2011.652746.

[15] Kai Yang and Basem El-Haik. *Design for six sigma : a roadmap for product development*. McGraw-Hill, New York, 2nd edition, 2009. ISBN 9780071547673 (alk. paper) 0071547673 (alk. paper).

[16] John M. Usher, Utpal Roy, and H. R. Parsaei. *Integrated product and process development : methods, tools, and technologies*. Wiley series in engineering design and automation. Wiley, New York, 1998. ISBN 0471155977 (cloth alk. paper).

[17] Thomas D. Hedberg Jr, Nathan W. Hartman, Phil Rosche, and Kevin Fischer. Identified research directions for using manufacturing knowledge earlier in the product life cycle. *International Journal of Production Research*, 55(3):819–827, 2017. ISSN 0020-7543. doi: 10.1080/00207543.2016.1213453.

[18] Technology. *Entry 4b*. Oxford University Press, Oxford, online edition, 2015.

[19] Innovation. *Entry 1a*. Oxford University Press, Oxford, online edition, 2015.

[20] Management. *Entry 1a*. Oxford University Press, Oxford, online edition, 2015.

[21] Colin Piddington and Mark Pegram. An ims test case - global manufacturing. In *Proceedings of the IFIP TC5/WG5.7 Fifth International Conference on Advances in Production Management Systems*, pages 11–20, 717937, 1993. North-Holland Publishing Co.

[22] Klaus-Dieter Thoben, Stefan A. Wiesner, and Thorsten Wuest. Industrie 4.0 and smart manufacturing – a review of research issues and application examples. *International Journal of Automation Technology*, 11(1):4–16, 2017. doi: 10.20965/ijat.2017.p0004.

[23] Joseph Tidd and J. R. Bessant. *Managing innovation : integrating technological, market, and organizational change*. Wiley, Hoboken, NJ, 4th edition, 2009. ISBN 9780470998106 (pbk. alk. paper).

[24] Logistics Information Service. Military engineering data asset locator system (medals). Report, Defense Logistics Agency, 2014. URL https://www.dlis.dla.mil/medals/pdfs/MEDALSHomepageOverview.pdf.

[25] Telecommunication Standardization Sector of ITU. Information technology – open systems interconnection – the directory – part 8: Public-key and attribute certificate frameworks, 2014. URL http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=64854.

[26] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 239: Application protocol: Product life cycle support, 2012.

Chapter 2

Literature Review

2.1 Introduction

A literature review was completed on two areas of research, digital thread and digital twin. Identifying current definitions, challenges, and trends for the two research areas was the goal of this review to justify the objectives of this proposed research. The main question is what are the current definitions, challenges, and trends, limited to the last 10 years, related to the *Digital Thread* and *Digital Twin* in the discrete manufacturing domain? Digital twin, also called digital surrogate, was included because its goal is to link the cyber space to the physical space for a given part. The digital thread enables the digital twin.

2.2 Systematic Literature Review Protocol

Developing a protocol is the first step in conducting a systematic literature review (SLR). Table 2.1 contains the information on the protocol for conducting the preliminary SLR of the Digital Thread and Digital Twin research areas. The protocol was developed in accordance to the National Institutes of Health recommendations [1] for conducting systematic reviews.

$$\text{Score} = 20 \times C_{TITLE} + 10 \times C_{ABS} + 5 \times C_{ASK}, \tag{2.1}$$

Where:

C_{TITLE} is the number of protocol keyword occurrences in the title of the paper,

C_{ABS} is the number of protocol keyword occurrences in the abstract of the paper, and

C_{ASK} is the number of protocol keyword occurrences in the author-selected keywords

2.3 Data Extraction

Four popular web-based publication search engines (ACM Digital Library [2], Engineering Village [3], IEEE Xplore [4], and Web of Science [5]) were used in the review because they

Table 2.1: A protocol for a Qualitative Systematic Review of Digital Thread / Digital Twin

Title:	Qualitative Systematic Review of Digital Thread and Digital Twin
Researchers:	Thomas Hedberg, Jr.;
Description:	This is a preliminary systematic literature review for two areas of research, digital thread and digital twin. Identifying current definitions, challenges, and trends for the two research areas is the goal of this review.
Objectives:	1. To identify current definitions, challenges, and trends related to the Digital Thread 2. To identify current definitions, challenges, and trends related to the Digital Twin
Main Question:	What are the current definitions, challenges, and trends, limited to the last 10 years, related to the Digital Thread and Digital Twin in the discrete manufacturing domain?
Keywords:	digital thread OR digital twin; AND challenges OR computer aided design OR cyber-physical systems OR decision making OR design OR gaps OR information management OR inspection OR life cycle OR lifecycle OR manufacture OR manufacturing OR product design OR quality assurance OR state of the art OR systems engineering OR trends OR uncertainty analysis
Source Selection Criteria:	known sources supporting indexing of engineering-related works
Studies Languages:	English;
Source Search Methods:	Web search through publication indexes;
Source Engine:	Engineering Village; Web of Science; Springer; IEEE; ACM;
Studies inclusion and exclusion criteria:	(I) high score; (I) survey paper; (I) definitions; (I) gaps, challenges, trends; (I) opportunities, needs; (I) data; (E) low score; (E) not related to topic; (E) full-text not found;
Studies types definition:	All types of studies will be considered;
Initial studies selection:	Based on the abstract; keywords and title; the inclusion and exclusion criteria will be applied;
Studies quality evaluation:	review of full paper
Information Extraction Fields:	type of study=[evaluation study, validation study, proposal of solution, results of experiment]; key takeaways;
Results Summarization:	literature has formed general consensus on the definition of "digital thread" and "digital twin"; various gaps and challenges are documented; several opportunities are presented; limited needs are described

are known to query indexes that include engineering-related publications. 139 papers were discovered during the literature search, of which 30 paper went through a full-text review. 23 papers were eventually selected for information extraction. An overview of the information extracted from the literature review is presented in Table 2.2.

Table 2.2: Overview of the information extracted from a full-text review of selected-literature related to Digital Thread and/or Digital Twin

Citation	**Year**	**Description**
Bajaj et al. [6]	2016	Provides an evaluative study of integrating system models with mechanical design. Also identifies opportunities and needs, while recommending solutions to gaps, challenges, and trends.
Adhikari et al. [7]	2016	Identifies gaps, challenges, and trends related to trust issues in big data about high-value manufactured parts.
Barnard Feeney et al. [8]	2015	Provides an overview and study of tolerancing standards integrated into the ISO 10303-242 (STEP AP242) standard. Also, identifies opportunities on how STEP AP242 may enable smart manufacturing.
Schrage [9]	2014	Provides recommendations for a systems approach to digital manufacturing and design innovation.
Kraft [10]	2016	Proposes definitions for digital thread and digital Twin. Also highlights life cycle integration opportunities and the use of computational and experimental knowledge.
Kraft [11]	2015	Identifies gaps, challenges, and trends in defining the concept of digital thread.
Hedberg Jr et al. [12]	2016	Provides validation and empirical data related to the benefits and return-on-investment for using model-based design, manufacturing, and inspection processes instead of paper-based processes.
Holzwarth et al. [13]	2012	Proposes a definition for digital thread as it relates to the U.S. Air Forces' requirements.
Schluse and Rossmann [14]	2016	Using the results of a simulation demonstration, provides recommended solutions and opportunities for digital twin.
Hochhalter et al. [15]	2014	Describes an evaluative study and provides opportunities for using digital twin to monitor damage in materials.
Scott-Emuakpor et al. [16]	2014	Proposes a definition for digital twin and provides opportunities for using a digital twin to infer material properties in nickel alloys.
Tuegel [17]	2012	Describes gaps and challenges for digital twin identified from an experiment related to airframes.
Wu et al. [18]	2016	Presents identified gaps, challenges, and trends in digital thread using a survey of cloud-based design and engineering analysis software. Tools

Table continued on next page

Table 2.2 – *Continued from previous page*

Citation	Year	Description
Canedo and ACM [19]	2016	Proposes a method to enable industrial-internet-of-things (IIoT) via a digital twin.
Beckmann et al. [20]	2016	Suggests opportunities for digital thread and digital twin using the Digital Manufacturing Commons (DMC) from the Digital Manufacturing and Design Innovation Institute (DMDII).
Rosen et al. [21]	2015	Describes gaps, challenged, and opportunities for the future of using digital twin in manufacturing.
Helu and Hedberg Jr [22]	2015	Identifies gaps and challenges to enabling smart manufacturing research and proposes using a product lifecycle test bed to enable taking advantage of the opportunities in digital thread.
Cerrone et al. [23]	2014	Proposes a definition for digital twin and describes an investigating on the effects of modeling as-manufactured geometry using digital twin.
Goto et al. [24]	2016	Presents opportunities and needs for the internet of things and describes the value for mechanical engineers in evolving commercial product lifecycle management systems.
Lu et al. [25]	2015	Provides a brief overview of the standards landscape and trends for smart manufacturing systems.
Alam and Saddik [26]	2017	Proposes a digital twin solution for cloud-based cyber-physical systems
Brodsky et al. [27]	2016	Identifies gaps, challenges, and opportunities in defining a system and architecture for reusable abstractions of manufacturing processes using the digital thread.
Wardhani and Xu [28]	2016	Proposes a definition of digital thread for model-based manufacturing based on the ISO 10303-242 (STEP AP242) standard.

2.4 Data Synthesis: Research Challenges, Opportunities, and Needs

Earlier, I identified three research objectives: (1) link data from the lifecycle phases using semantic-web concepts, (2) ensure authentication, authorization, and traceability of product data, and (3) integrate feed-back and feed-forward process control of product lifecycle processes using data-driven methods. These research objectives align with the results of the my literature review described in Table 2.2. Based on the review results, the research objectives would output novel contributions to the manufacturing domain.

The following list classifies the output from the previous works selected from the preliminary literature review. The output of the review is classified according to the three proposed research objectives. Each research objective is listed as a parent node (bold typeface) in the ordered-list below. The literature review output is classified under the corresponding research objective that addresses the challenge, opportunity, and/or need for that research objective.

1. **Link data across lifecycle phases using semantic-web concepts**
   (a) Challenges
      i. Standard information representations are lacking and/or are difficult to manage [7]
      ii. Selecting and integrating sub-models from various and across domains (i.e., model interoperability) [17]
      iii. Ensuring the proper system behavior to reach the desired goals is increasing in complexity [21]
      iv. Variety in domain-specific tools and in stakeholder roles and perspectives [27]
   (b) Opportunities
      i. Mirroring cyber objects with physical objects would resolve differences using monitoring, diagnostics, and prognostics methods [26]
      ii. Increased creation, use, and structured storage of digital artifacts that are needed by different stakeholders where all elements are connected and there exists meta-information as well as semantics [21]
   (c) Needs
      i. Means of gathering, storing, and processing all data in the lifecycle [21]
      ii. Quick means of discovering and retrieving data when and where it is needed [21]
      iii. Access to realistic models of current states of processes and their behavior in interacting with the real-world environment [21]
      iv. Integration of new manufacturing-focused models with existing computer-aided technologies [27]
2. **Ensure authentication, authorization, and traceability of product data**
   (a) Challenges
      i. Managing provenance information is done manually and is burdensome [7]
      ii. Data trustworthiness, access control, and confidentiality remains a high-priority [7]
   (b) Opportunities
      i. Monitoring will enable detection of unforeseen events in near-real-time and provide the digital twin with updates of actual usage and states [15]

(c) Needs

i. (No needs were identified in the preliminary literature review. However, needs likely exist given the challenges and opportunities identified.)

3. **Implement feed-back and feed-forward process control to decisions**

(a) Challenges

i. Reducing cycle time during development, test and evaluation (DT&E) [11]

ii. Managing and reducing the uncertainty of models and decisions across multiple domains [17]

iii. Ensuring the proper system behavior to reach the desired goals is increasing in complexity [21]

(b) Opportunities

i. Feedback loops that are always active and improving the quality of service of the physical systems [26]

ii. Enable virtual prototyping and experimentation earlier in the lifecycle using data collected from the lifecycle [11]

iii. Reduce uncertainty across the lifecycle by incorporating into available models as much initial and in-service information as possible [15]

(c) Needs

i. Next-generation software and computing architectures to effectively mine data and use it to solve complex problems [27]

ii. Automate feedback of later lifecycle events to earlier phases in the lifecycle [15]

iii. Enable decision-making based on a wide range of technical and business parameters [27]

The literature review provides evidence to support that the three proposed research objectives are valid and novel. Digital twin has received considerably more attention in the literature. However, the digital twin requires a digital thread to be successful.

Moreover, both digital thread and digital twin are emerging areas of research. This is evident from the fact that most of the literature was published in the last two years and very little archival journal articles are available. The majority of previous work has been presented in conferences. This supports the fact that digital thread and digital twin concepts are still emerging.

2.5 Required knowledge

Several areas of knowledge must be integrated to complete the proposed research. These areas include, but are not limited to: open data, semantic web, model-based enterprise (MBE), and data authentication / authorization / traceability. The remainder of this section is a very brief description of the literature in these areas. More in-depth review of these areas are provided in later chapters.

An open-data culture is growing globally. Technology is enabling the open-data revolution – governments, academic institutions, and industries are using data to create knowledge about the world and make decisions [29]. Open Knowledge Foundation [29] defines open data as, "data and content that can be freely used, modified, and shared by anyone for any purpose [29]." There are several examples [30, 31, 32] of open data in the sciences. Open data is lacking from the manufacturing domain, but some data-linking examples [33, 34] do exist.

The Semantic Web functions through machine-interpretable access to structured collections of information [35]. The first step in building the Semantic Web was to start building higher-order relationships between data. The concept of *Linked Data* provides a foundation for defining high-order relationships between sets of data. Berners-Lee [36] proposed four rules for Linked Data, which are: (1) use Uniform Resource Identifiers (URIs) [37] as names for things, (2) use Hypertext Transfer Protocol (HTTP) URIs so names are discoverable, (3) provide useful information using standards (e.g., Resource Description Framework (RDF) [38], SPARQL [39]), and (4) include links to other URIs so n-order links are discoverable.

MBE has introduced new requirements on data usage in manufacturing systems. MBE calls for each phase and function of the product lifecycle to adopt model-based data standards to effectively integrate data for efficient reuse and exchange between product lifecycle phases. The need for automated methods to collect, transmit, analyze, and act on the most appropriate data is gaining attention [22, 40, 41, 42].

From the perspective of product design and manufacturing, traceability is defined as the ability to discover the history of decisions in the lifecycle, control the quality of data, products, and processes, and understand the relationship between assets [43, 44, 45, 46, 47]. Tracing dependency links between assets supports establishing relationships between those assets [44]. Understanding those links and relationships helps determine how decisions made during the creation and modification of assets affect related assets. Therefore, traceability may be considered a critical quality attribute intended to ensure system outputs conform to stakeholder requirements [45, 46]. Data traceability cannot be separated from authentication and authorization. Authentication is the act of determining that an entity (e.g., person, data) is as the entity is declared and authorization is the process of determining what permissions an entity is granted by a trusted source.

Chapter Bibliography

[1] NIH Library. Systematic reviews: Systematic review protocols and protocol registries, 2015. URL http://nihlibrary.campusguides.com/c.php?g=38332&p=244525.

[2] Association for Computing Machinery. ACM Digital Library, 2017. URL http://dl.acm.org/.

[3] Elsevier B.V. Engineering Village, 2015. URL https://www.engineeringvillage.com.

[4] IEEE. IEEE Xplore, 2017. URL http://ieeexplore.ieee.org.

[5] Thomason Reuters. Web of Science, 2017. URL http://www.webofknowledge.com.

[6] Manas Bajaj, Dirk Zwemer, and Bjorn Cole. Architecture to geometry - integrating system models with mechanical design. In *AIAA Space and Astronautics Forum and Exposition, SPACE 2016, September 13, 2016 - September 16, 2016*, AIAA Space and Astronautics Forum and Exposition, SPACE 2016. American Institute of Aeronautics and Astronautics Inc, AIAA.

[7] Anku Adhikari, Avesta Hojjati, Juanli Shen, Jui-Ting Hsu, William P. King, and Marianne Winslett. Trust issues for big data about high-value manufactured parts. In *2nd IEEE International Conference on Big Data Security on Cloud, IEEE BigDataSecurity 2016, 2nd IEEE International Conference on High Performance and Smart Computing, IEEE HPSC 2016 and IEEE International Conference on Intelligent Data and Security, IEEE IDS 2016, April 9, 2016 - April 10, 2016*, Proceedings - 2nd IEEE International Conference on Big Data Security on Cloud, IEEE BigDataSecurity 2016, 2nd IEEE International Conference on High Performance and Smart Computing, IEEE HPSC 2016 and IEEE International Conference on Intelligent Data and Security, IEEE IDS 2016, pages 24–29. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/BigDataSecurity-HPSC-IDS.2016.50.

[8] Allison Barnard Feeney, Simon P. Frechette, and Vijay Srinivasan. A portrait of an ISO STEP tolerancing standard as an enabler of smart manufacturing systems. 15, 2015. ISSN 15309827. doi: 10.1115/1.4029050. URL http://dx.doi.org/10.1115/1.4029050.

[9] Dan Schrage. Providing a systems approach for digital manufacturing and design innovation. In *70th American Helicopter Society International Annual Forum 2014, May 20, 2014 - May 22, 2014*, volume 4 of *Annual Forum Proceedings - AHS International*, pages 3213–3227. American Helicopter Society. ISBN 15522938.

[10] Edward M. Kraft. The US Air Force digital thread/digital twin life cycle integration and use of computational and experimental knowledge. In *54th AIAA Aerospace Sciences Meeting, 2016, January 4, 2016 - January 8, 2016*. American Institute of Aeronautics and Astronautics, 2016.

[11] E. M. Kraft. Hpcmp createtm-av and the air force digital thread. In *53rd AIAA Aerospace Sciences Meeting, 5-9 Jan. 2015*, 53rd AIAA Aerospace Sciences Meeting, page 13 pp. American Institute of Aeronautics and Astronautics.

[12] Thomas D. Hedberg Jr, Joshua Lubell, Lyle Fischer, Larry Maggiano, and Allison Barnard Feeney. Testing the digital thread in support of model-based manufacturing and inspection. *Journal of Computing and Information Science in Engineering*, 16(2): 1–10, 2016. ISSN 1530-9827. doi: 10.1115/1.4032697.

[13] Richard Holzwarth, Eric Tuegel, and Pam Kobryn. Airframe digital twin: Creating virtual replicas of every aircraft in the fleet. In *Prognostics and Health Management Solutions Conference - PHM: Driving Efficient Operations and Maintenance, MFPT 2012, April 24, 2012 - April 26, 2012*, Technical Program for MFPT 2012, The Prognostics and Health Management Solutions Conference - PHM: Driving Efficient Operations and Maintenance. Machine Failure Prevention Technology Society (MFPT).

[14] Michael Schluse and Juergen Rossmann. From simulation to experimentable digital twins: Simulation-based development and operation of complex technical systems. In *2nd Annual IEEE International Symposium on Systems Engineering, ISSE 2016, October 3, 2016 - October 5, 2016*, ISSE 2016 - 2016 International Symposium on Systems Engineering - Proceedings Papers, page IEEE; IEEE Systems Council. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/SysEng.2016.7753162.

[15] J. Hochhalter, P. W. Leser, A. J. Newman, K. V. Gupta, V. Yamakov, R. S. Cornell, A. S. Willard, and G. Heber. Coupling damage-sensing particles to the digitial twin concept. Report nasa/tm-2014-218257; l-20401, 2014.

[16] Onome Scott-Emuakpor, Tommy George, Joseph Beck, Jeremy Schwartz, Casey Holycross, M. H. Herman Shen, and Joseph Slater. Material property determination of vibration fatigued dmls and cold-rolled nickel alloys. In *ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, GT 2014, June 16, 2014 - June 20, 2014*, volume 7A of *Proceedings of the ASME Turbo Expo*, page International Gas Turbine Institute. American Society of Mechanical Engineers (ASME). doi: 10.1115/GT2014-26247.

[17] Eric J. Tuegel. The airframe digital twin: Some challenges to realization. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 2012, April 23, 2012 - April 26, 2012*, 53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 2012, page American Institute of Aeronautics and Astronautics (AIAA). American Institute of Aeronautics and Astronautics Inc.

[18] D. Z. Wu, J. Terpenny, and D. Schaefer. A survey of cloud-based design and engineering analysis software tools. In *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, 2016, Vol 1a*. Amer Soc Mechanical Engineers. ISBN 978-0-7918-5007-7.

[19] A. Canedo and ACM. Industrial IoT lifecycle via digital twins. *2016 International Conference on Hardware/Software Codesign and System Synthesis (Codes+Isss)*, page 1, 2016. doi: 10.1145/2968456.2974007.

[20] B. Beckmann, A. Giani, J. Carbone, P. Koudal, J. Salvo, and J. Barkley. *Developing the Digital Manufacturing Commons: A National Initiative for US Manufacturing Innovation*, volume 5 of *Procedia Manufacturing*, pages 182–194. Elsevier Science Bv, Amsterdam, 2016. doi: 10.1016/j.promfg.2016.08.017.

[21] R. Rosen, G. von Wichert, G. Lo, and K. D. Bettenhausen. About the importance of autonomy and digital twins for the future of manufacturing. *Ifac Papersonline*, 48(3): 567–572, 2015. ISSN 2405-8963. doi: 10.1016/j.ifacol.2015.06.141.

[22] Moneer Helu and Thomas Hedberg Jr. Enabling smart manufacturing research and development using a product lifecycle test bed. *Procedia Manufacturing*, 1:86–97, 2015. ISSN 2351-9789. doi: http://dx.doi.org/10.1016/j.promfg.2015.09.066.

[23] A. Cerrone, J. Hochhalter, G. Heber, and A. Ingraffea. On the effects of modeling as-manufactured geometry: Toward digital twin. *International Journal of Aerospace Engineering*, page 10, 2014. ISSN 1687-5966. doi: 10.1155/2014/439278.

[24] S. Goto, O. Yoshie, and S. Fujimura. Internet of things value for mechanical engineers and evolving commercial product lifecycle management system. In *2016 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 1021–1024. doi: 10.1109/IEEM.2016.7798032.

[25] Y. Lu, K. C. Morris, and S. Frechette. Standards landscape and directions for smart manufacturing systems. In *2015 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 998–1005. ISBN 2161-8070. doi: 10.1109/CoASE.2015.7294229.

[26] K. M. Alam and A. El Saddik. C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access*, PP(99):1–1, 2017. ISSN 2169-3536. doi: 10.1109/ACCESS.2017.2657006.

[27] A. Brodsky, M. Krishnamoorthy, W. Z. Bernstein, and M. O. Nachawati. A system and architecture for reusable abstractions of manufacturing processes. In *2016 IEEE International Conference on Big Data (Big Data)*, pages 2004–2013. doi: 10.1109/BigData.2016.7840823.

[28] R. Wardhani and X. Xu. Model-based manufacturing based on STEP AP242. In *2016 12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pages 1–5, 2016. doi: 10.1109/MESA.2016.7587187.

[29] Open Knowledge Foundation. About, 2015. URL https://okfn.org/about/.

[30] National Aeronautics and Space Administration. Physical science information: History and purpose, 2014. URL http://psi.nasa.gov/History.aspx.

[31] Consortium for Ocean Leadership. Final network design. Report, Ocean Observatories Initiative, 2010. URL oceanleadership.org/wp-acontent/uploads/2009/04/1101-00000_FND_OOI_2010-04-22_ver_2-06_public1.pdf.

[32] EarthCube. Earthcube strategic science plan. Report, National Science Foundation, 2015. URL http://earthcube.org/document/earthcube-strategic-science-plan.

[33] Logistics Information Service. Military engineering data asset locator system (MEDALS). Report, Defense Logistics Agency, 2014. URL https://www.dlis.dla.mil/medals/pdfs/MEDALSHomepageOverview.pdf.

[34] Mark R. Cutkosky, Jay M. Tenenbaum, and Jay Glicksman. Madefast: collaborative engineering over the internet. *Communications of the ACM*, 39(9):78–87, 1996. ISSN 0001-0782. doi: 10.1145/234215.234474. URL http://dl.acm.org/citation.cfm?id=234474.

[35] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific American*, May, 2001. URL http://www.scientificamerican.com/article/the-semantic-web/.

[36] Tim Berners-Lee. Linked data, 2009 2006. URL http://www.w3.org/DesignIssues/LinkedData.html.

[37] World Wide Web Consortium. Semantic web, 2006. URL https://www.w3.org/standards/semanticweb/.

[38] World Wide Web Consortium. Resource description framework (RDF), 2014. URL http://www.w3.org/RDF/.

[39] World Wide Web Consortium. SPARQL 1.1 overview, 2013. URL http://www.w3.org/TR/sparql11-overview/.

[40] Energetics Inc. Measurement science roadmap for prognostics and health management for smart manufacturing system. Report, National Institute of Standards and Technology, 2015.

[41] R. Gao, L. Wang, R. Teti, D. Dornfeld, S. Kumara, M. Mori, and M. Helu. Cloud-enabled prognosis for manufacturing. *CIRP Annals - Manufacturing Technology*, 64(2): 749–772, 2015. ISSN 0007-8506. doi: 10.1016/j.cirp.2015.05.011. URL http://www.sciencedirect.com/science/article/pii/S000785061500150X.

[42] Min Li, Shuming Gao, and Charlie C. Wang. Real-time collaborative design with heterogeneous CAD systems based on neutral modeling commands. *Journal of Computing and Information Science in Engineering*, 7(2):113–125, 2006. ISSN 1530-9827. doi: 10.1115/1.2720880. URL http://dx.doi.org/10.1115/1.2720880.

[43] V. L. Hamilton and M. L. Beeby. Issues of traceability in integrating tools. In *IEEE Colloquium on Tools and Techniques for Maintaining Traceability During Design*, pages 4/1–4/3, 1991.

[44] B. Ramesh. Process knowledge management with traceability. *IEEE Transactions on Software Engineering*, 19(3):50–52, 2002. ISSN 0740-7459. doi: 10.1109/MS.2002.1003454.

[45] Kannan Mohan and Balasubramaniam Ramesh. Traceability-based knowledge integration in group decision and negotiation activities. *Decision Support Systems*, 43(3): 968–989, 2007. ISSN 0167-9236. doi: http://dx.doi.org/10.1016/j.dss.2005.05.026. URL http://www.sciencedirect.com/science/article/pii/S0167923605000916.

[46] Kannan Mohan, Peng Xu, Lan Cao, and Balasubramaniam Ramesh. Improving change management in software development: Integrating traceability and software configuration management. *Decision Support Systems*, 45(4):922–936, 2008. ISSN 0167-9236. doi: http://dx.doi.org/10.1016/j.dss.2008.03.003. URL http://www.sciencedirect.com/science/article/pii/S0167923608000523.

[47] M. Z. Ouertani, S. Baïna, L. Gzara, and G. Morel. Traceability and management of dispersed product knowledge during design and manufacturing. *Computer-Aided Design*, 43(5):546–562, 2011. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/j.cad.2010.03.006. URL http://www.sciencedirect.com/science/article/pii/S0010448510000618.

Chapter 3

Background Knowledge

3.1 Data and Information

The concept of linking cross-domain data is not new. The contribution of the Lifecycle Information Framework and Technology (LIFT) concept is the implementation and integration of multiple existing technologies, paradigms, and concepts in a novel way to form a framework that would provide significant benefit to manufacturing industry groups. In this section, I provide background information about concepts and technologies that have informed the development of the LIFT concept. In each sub-section, a review is provided on similar efforts being deployed and/or tested in scientific domains and discuss how the similar concepts informed the work.

3.1.1 Open Data

An open-data culture is growing globally. Technology is enabling the open-data revolution – governments, academic institutions, and industries are using data to create knowledge about the world and make decisions [1]. In 2005, the Open Knowledge Foundation [1] provided the first definition of "open data." They define it as, "open data and content that can be freely used, modified, and shared by anyone for any purpose [1]." Further, the Open Knowledge Foundation [1] believes, "open knowledge can empower everyone, enabling people to work together to tackle local and global challenges, understand our world, expose inefficiency and challenge inequality and hold governments and companies to account."

In 2006, the Open Knowledge Foundation launched CKAN [2]. CKAN is an open-source data portal platform to streamline publishing, sharing, finding, and using data [3]. CKAN supports government open-data initiatives in several countries [2], including Austria, Brazil, Canada, Germany, Netherlands, Norway, the United Kingdom, the United States, and Uruguay. CKAN provides the infrastructure sufficient for sharing data across the globe, but it lacks the ability on its own to provide trust and traceability in the data. The Open Data Institute provides a solution to fill the trust and traceability gap.

The Open Data Institute (ODI), founded by Tim Berners-Lee and Nigel Shadbolt, is a non-profit organization looking to accelerate the evolution of the open-data culture [4]. The UK Government provided the ODI with £10 million in seed funding. The ODI strives to create economic, environmental, and social value through open data. The ODI is developing the "open data certificate" to help data publishers explain what the open data is about and deliver open data that people can trust [5]. The ODI certificate is similar to the X.509-based [6] digital certificates used widely across the Internet today. The goal of the ODI certificate is to provide sufficient meta-data about the open data so users can vet the fidelity of the data [5]. The ODI provides four certificate types [7]:

1. Raw: basic meta-data for publishing open data with no support or additional guarantees
2. Pilot: open data users receive extra support from, and can provide feedback to the publisher.
3. Standard: regularly published open data with robust support that people can rely on
4. Expert: a complete information infrastructure – the ODI says no one has achieved this level of certification

The ODI certificates may act as a "seal of approval" to build trust in open data. A framework that includes the CKAN and ODI certificate technologies delivers significant impact to the public. Government open data initiatives, like the United States' Data.gov [8], demonstrate the impact of open data through increasing the transparency and accessibility of data that the public (e.g., researchers, journalists, students) can use. The U.S. General Services Administration [9] claims that "open government data is important because the more accessible, discoverable, and usable data is the more impact it can have. These impacts include, but are not limited to: cost savings, efficiency, fuel for business, improved civic services, informed policy, performance planning, research and scientific discoveries, transparency and accountability, and increased public participation in the democratic dialogue." Data.gov provides access to federal, state, and local data to support activities such as conducting research (e.g., Smart Grid studies), developing applications (e.g., mobile apps), and designing data visualizations (e.g., geo-spatial analysis from geographic data) [9].

3.1.2 Open Data Initiatives in the Sciences

The expansion of the open-data culture is enabling scientific communities to develop a deep understanding needed to reform data collection and sharing in the sciences. Three successful applications of open-data concepts are the Physical Science Informatics (PSI) system [10], the Ocean Observatories Initiative (OOI) [11], and EarthCube [12]. These examples show how open data can accelerate scientific discovery through a "force multiplier" effect gained through the ability to reuse or expand upon shared datasets. Further, the PSI, OOI, and EarthCube examples may provide the foundation for the eventual development of a data observatory.

Scientists have been conducting micro-gravity experiments on the International Space Station (ISS) since 2001. The National Aeronautics and Space Administration (NASA) developed a repository to store all of the raw or minimally processed data from scientific experiments conducted on the ISS. The NASA PSI system [10] makes the stored data available to the public. A Director of Space Life and Physical Sciences at NASA stated that [10] "[open data] brings together the community of researchers to define an envelope of experiments that will be conducted and analyzed, leveraging modern high content analytics in the life and physical sciences. The resulting data from that envelope of experiments will then be used to create experimental informatics libraries that will support many more investigators and funded ISS-derived research. What that does is, it converts what would be normally a single principal investigator (PI) research opportunity into multiple PI research opportunities now and into the future."

Where NASA's PSI system provides data curation for a single experiment type, the OOI [11] and EarthCube [12] provide data curation for complex sets of experiment types. OOI determined studying the complex interaction of biological, chemical, physical, and geological processes in the ocean is limited severely by available technical infrastructures [11]. In response, the OOI developed an interactive cyber-infrastructure with common design elements to enable integrating multiple scales of ocean observations [11]. The OOI cyber-infrastructure brings researchers together for data curation, discovery, and sharing of experimental data captured throughout the OOI system. The OOI infrastructure provides persistence, adaptability, interoperability, and community related to the data in the system [11].

EarthCube [12] uses a concept similar to the OOI for monitoring geo-sciences data. In 2011, the National Science Foundation (NSF) posed a, "challenge for the many academic, agency, and industry stakeholders in the geo-, cyber-infrastructure, computer, and social sciences to create new capabilities for sharing data and knowledge and conducting research" [12]. EarthCube discovered five common themes that called for the integration of data across different domains. EarthCube recognizes that the complexity of different domain models presents a key challenge to the geo-sciences community. But EarthCube believes, "to cultivate future generation of researchers and also for EarthCube to be of use to policy and decision makers, geo-scientists must be able to retain access to and communicate the results and uncertainties of information and research that advances knowledge to society" [2015].

3.1.3 Manufacturing Data in the U.S. Department of Defense

The Department of Defense Authorization Act, 1985 [13, Sec. 1252] created a mandate upon the U.S. U.S. Department of Defense (DoD) to develop a centralized system for the management of technical data related to items of supply. In 1985, the Office of the Secretary of Defense directed the Defense Logistics Agency (DLA) to establish and operate the centralized system. In 1988, DLA implemented the Military Engineering Data Asset Locator

System (MEDALS). MEDALS is a globally accessible networked system where engineering drawings and technical documents reside [14]. DLA [14] describes the system as a research tool to act as a "first-discovery mechanism" to assist a user with finding engineering documents when the user does not know where the document might reside. MEDALS also houses the specific repository in which the documents reside.

As of 2014, MEDALS tracked the location of 44 million engineering data assets across 45 different data repositories [14]. However, use of MEDALS has decreased recently because a significant amount of human input is required to keep the system up to date. DLA recognized MEDALS needed a web service interface to enable automating data-asset management. DLA subsequently added Extensible Markup Language (XML) ingestion capabilities and enhanced batch querying capabilities to support a MEDALS web service interface [14].

The DoD also supported and still supports collaborative design research and development. The Defense Advanced Research Projects Agency (DARPA) Manufacturing Automation and Design Engineering (MADE) program [15] was one of the first collaborative design research projects to use a digital infrastructure. The goal of MADE was to develop Internet-based tools, services, protocols, and design methodologies to support the design activities of geographically distributed expert teams [15]. The project tested the ability of a diverse and geographically dispersed team to design a product collaboratively. The MADE program recognized the important need to curate the design process as much as the actual design. The project team developed the "Madefast Web," which was a set of web pages, shared via the World Wide Web (WWW), that acted as a repository for the computer-aided design (CAD) models, notes, test results, calculations, and other information relating to the design [15]. At the top level of the Madefast Web was an index that created links to all of the MADE project participant-hosted pages. This allowed the design authorities for the product sub-components to maintain control of the needed data, but still share it to the rest of the MADE project team for reuse and re-purposing. Because the MADE tools are Internet-based, information and applications were simply accessed through the point and click of a mouse.

The MADE program provides an interesting example of a knowledge base and/or expert system. Cutkosky et al. [15] said, "Madefast uses the WWW as a corporate memory, sharing design information across the design team, and preserving it for downstream tasks such as maintenance and redesign. Such information is, of course, useless if it cannot be found. The standard approaches to locating information on WWW pages, such as hierarchical directories and keyword searches, do not provide adequate granularity and precision for engineering design applications." In some aspects, the MADE project was ahead of its time. The benefits of the Madefast concept became evident when the project team decided to make a second version of the design product. The project team was able to review all of the design process and product documentation of the first design while on conference calls. The design information was readily available to accelerate decisions for the new design version.

The MADE program also identified a few key issues that still exist today. Data organization within the Madefast Web was difficult. Although three major overhauls were made in an effort to keep data organized, newcomers remained unable to find information quickly [15]. Human interaction and different workplace cultures across the team also remained a challenge. However, the most significant issue in the MADE project related to systems and services integration.

The MADE program identified several technological advances that were needed before integrated services would be able to compete with current practices. The issues were related to security, standards, and interoperability. The MADE program highlighted project team members' data security and intellectual property concerns. In addition, the MADE program determined that standards are an important enabler for using house-made tools with any service provider or partner [15]. Interestingly, all of the challenges identified two decades ago by the MADE program persist in industry today.

3.1.4 Semantic Web

The Oxford Dictionary [16] defines *semantic* as, "relating to meaning in language or logic." In the early 2000's, several papers and magazines [17, 18, 19] wrote about adding semantic definition to the WWW and developing the concept of the *Semantic Web*. The World Wide Web Consortium (W3C) declares the Semantic Web to be the integration of linked data, vocabularies, query, inference, and vertical applications [20]. The WWW started out as a cluster of documents published via the Internet and linked to other documents using hypermedia methods [18]. Markup languages (e.g., Hypertext Markup Language (HTML)) are used to deliver content to people in human-readable form. The traditional document-centric WWW provides little focus on the representation of the data contained in the documents [18]. The purpose of the Semantic Web is to provide machine-readable data that represents the semantics of the concepts presented in traditional WWW documents. Tim Berners-Lee [17] – the recognized father of the WWW – stated that "the Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation."

The Semantic Web functions through machine-interpretable access to structured collections of information [17]. The first step in building the Semantic Web was to start building higher-order relationships between data. The concept of *Linked Data* provides a foundation for defining high-order relationships between sets of data. Berners-Lee [19] proposed four rules for Linked Data, which are: (1) use Uniform Resource Identifiers (URIs) [20] as names for things, (2) use Hypertext Transfer Protocol (HTTP) URIs so names are discoverable, (3) provide useful information using standards (e.g., Resource Description Framework (RDF) [21], SPARQL [22]), and (4) include links to other URIs so n-order links are discoverable.

However, linking data is simply not enough to enable the Semantic Web. It is important that the data is organized and sufficient structure and context exist to support the evolution from data to information to knowledge. Vocabularies, or *ontologies*, support the query and inference components of the Semantic Web. Ontologies improve the accuracy of WWW searches at the basic level [17]. In more complex systems (e.g., the Semantic Web), ontologies relate the information of one system to the associated knowledge structures and inference rules used to apply semantic representation using data [17].

In 2009, Khilwani et al. [23] presented a survey of 14 different ontologies relevant to the manufacturing domain. Since then [24, 25, 26] have discussed additional ontology studies. While having only one ontology to encapsulate the entire product lifecycle would be ideal, achieving consensus on the definition of that ontology is unlikely because of the many different viewpoints that exist across the lifecycle. Therefore, what matters more is the ability to link data through the product lifecycle and apply the necessary context to the data to ensure the right information and knowledge is available to the roles and functions when they need it. Enabling linked data by normalizing the process for linking different ontologies in the Semantic Web enables a powerful data curation and discovery mechanism. Thus, integrating the Semantic Web with the product lifecycle provides users (e.g., humans, machines) the ability to build queries for discovering complex information, which is paramount to discovering and extracting data for the user.

However, inference may be the most important component of the Semantic Web. Inference enables automatic analysis of data. The automatic analysis, built with reasoning and other artificial intelligence algorithms, supports managing knowledge in the Semantic Web. Further, inference provides automatic procedures to generate new relationships based on the data and on additional information from ontologies [27].

The Semantic Web, "provides a declarative interoperable foundation for modeling, encoding, dissemination, and process [different] types of knowledge in a common, interoperable way... [18]." Technologies like RDF [21], Ontology Web Language (OWL) [28], Simple Knowledge Organization System (SKOS) [29], and SPARQL [22] provide important tools and the foundation of the Semantic Web.

Several domains have implemented the Semantic Web as vertical applications within their communities. Examples of domains working within the Semantic Web are health care and life sciences, social spaces, digital libraries, financial services, oil and gas exploration, and e-Government [30]. The health care and life sciences domain has published extensively on the topic with examples such as [31, 32, 33, 34].

Previous work in Semantic Research [25, 35, 36] provide examples for the manufacturing domain. However, much of the Semantic Web research in manufacturing is limited to a single component of the Semantic Web such as inference systems or ontology development. There has been little investigation into fully integrating all of the Semantic Web components (i.e., linked data, vocabularies, query, inference, and vertical applications) into manufacturing and across the product lifecycle. Further, the ontologies discussed in [23, 24, 25, 26] do not address integrating cross-domain (e.g., design, manufacturing, quality) ontologies to generate an ontology for the entire product lifecycle. Design information may not be linked directly to quality information (e.g., [24]) and the manufacturing information may only be linked to a small portion of design information that is directly relevant to the manufacturing viewpoint (e.g., [25]). In other words, the Semantic Web in manufacturing remains as silos of domain-specific data linked only within the phase of the product lifecycle that generated the data. Therefore, the manufacturing domain remain an untapped opportunity for implementing the Semantic Web to provide significant impact with a full-integration of the product lifecycle.

3.1.5 Open Data Standards

Information technology advances (e.g., data analytics, service-oriented architectures, and networking) have triggered a digital revolution [37] that when coupled with operational technology (e.g., hardware and software for sensing, monitoring, and control of product and processes) holds promise for reducing costs, improving productivity, and increasing output quality. Modern manufacturing enterprises are both more globally distributed and digital, resulting in increasingly complex manufacturing system networks [36, 38]. Manufacturers are under mounting pressure to perform digital manufacturing more efficiently and effectively within these distributed-manufacturing systems. To do so, industry is changing how product definitions are communicated – from paper to models. This transition is being called model-based enterprise (MBE).

MBE has introduced new requirements on data usage in manufacturing systems. MBE calls for each phase and function of the product lifecycle to adopt model-based data standards to effectively integrate data for efficient reuse and exchange between product lifecycle phases. The need for automated methods to collect, transmit, analyze, and act on the most appropriate data is gaining attention [39, 40, 41, 42].

In addition, the MBE strategy must ensure model-based-data interoperability between design activities (e.g., product and assembly design) and manufacturing activities (e.g., fabrication, assembly, and quality assurance). ISO 10303-242:2014 [43], ISO 32000 [44] and ISO 14739 [45], ISO 6983 [46], MTConnect [47], and Quality Information Framework (QIF) [48] are three emerging standards that show promise for enabling linked data throughout the product lifecycle.

ISO 10303-242 (STEP AP242)

The standard, ISO 10303-242:2014 [43] titled "Managed Model Based 3D Engineering," commonly known as Standard for the Exchange of Product Model Data Application Protocol 242 (STEP AP242), is an international standard that shows promise for enabling linked data. The goal of STEP AP242 is to support a manufacturing enterprise with a range of standardized information models that flow through a long and wide "digital thread'' that makes the manufacturing systems in the enterprise smart [49]. Digital data plays a central role in achieving the goal of STEP AP242.

Published in December 2014, STEP AP242 contains extensions and significant updates to other Standard for the Exchange of Product Model Data (STEP) Application Protocols (APs) for product and manufacturing information (PMI), kinematics, and tessellation [50]. PMI is the presentation and representation of geometric dimensions and tolerances (GD&T), material specifications, component lists, process specifications, and inspection requirements within a three-dimensional (3D) product definition [51]. A study [51], comparing drawing-based processes to model-based processes, concluded that PMI has the potential to make many lifecycle processes run faster, with fewer errors, and at lower cost, since STEP AP242 offers standards-based models that include the representation of PMI that is computer interpretable [49]. This is a major breakthrough that supports manufacturing's need for model-based computer-aided manufacturing (CAM) and coordinate-measurement system (CMS) processes because STEP AP242 increases the effectiveness of MBE by enabling a common path for model-based definition (MBD) and model-based manufacturing (MBM) integration [50, 52].

ISO 32000 and ISO 14739 (PDF/PRC)

The Portable Document Format (PDF) [44] and Product Representation Compact (PRC) [45] international standards are often combined into technologies for visualizing product-definition data. PDF/PRC is often referred to as a 3D PDF. While there are several "flavors" of 3D PDF, the combination of PRC embedded in a PDF document is emerging as the industry recommended practice. PDF/PRC enables the display of 3D product definition in any PDF reading software that conforms to the standard. Using PDF/PRC enables effective and efficient visualization of product data throughout the lifecycle for human-consumption. For the concept proposed in this paper, PDF documents are generated from native CAD data and linked to the native data.

ISO 6983 (G-CODE)

ISO 6983-1:2009 [46] is an international standard that defines the data format to program position, line motion, and contouring control systems in the numerical control (NC) of machines. This data format is commonly known as G-code. G-code was created at MIT in the late 1950s and, like CAD, rose in popularity through the 1970s [53]. Today, G-code is the near-universal format for programming computer-based NC machines.

G-code is generated typically from a manufacturing plan using a CAM system. G-code files are defined using a standardized ASCII-based set of commands. Each line of the G-code is a new command to the machine. Header information is standardized to support some traceability. For the concept proposed in this paper, additional header information and metadata is added to the G-code to support linking the G-code back to both STEP and CAM data.

MTConnect

MTConnect is an open-source, read-only data-exchange standard for manufacturing equipment and applications developed by the Association for Manufacturing Technology (AMT) [47]. It is based on XML and HTTP and provides information models and communications protocols to enhance the data acquisition capabilities of manufacturing equipment and applications and to enable a plug-and-play environment. While other communication protocols may exist for data transfer, the information models defined in MTConnect are the only common vocabulary and structure created for manufacturing equipment data. Perhaps the most important type of data addressed by the standard is real and near-realtime data from the equipment (e.g., current speed or position, program blocks). This ability is critical in enabling the standard to support the digital thread by providing data and information on the as-built condition of a part.

The MTConnect standard contains four types of information models for manufacturing equipment: Devices, Streams, Assets, and Errors [47]. Devices provides the metadata that describes the hierarchical structure of a device (i.e., a piece of manufacturing equipment) and its components and available data items. Streams provides the actual data values returned by a piece of manufacturing equipment for each data item coupled with a time stamp. Assets provides additional information models to describe systems that are used by a piece of manufacturing equipment but are not part of the equipment or its components, e.g., cutting tools or peripheral systems that are part of the part flow, sequence, or process steps. Errors focuses on protocol-related errors.

Users access the data and information provided by the information models in MTConnect through the Agent [47]. The Agent implements the MTConnect protocol and generates the relevant XML document. MTConnect only standardizes communication between the Agent and an application. The Agent returns data only when requested by an application. It acts as an HTTP server and uses Representational State Transfer (REST) when interacting with any data source, In other words, it is the responsibility of the data source to send updates of its state to the Agent.

ANSI/DMSC Quality Information Framework

The QIF [48] is an American National Standards Institute (ANSI) standard sponsored by the Dimensional Metrology Standards Consortium (DMSC). QIF defines an integrated set of XML information models that enable the effective exchange of metrology data throughout the entire metrology process. QIF handles feature-based dimensional metrology, quality measurement planning, first article inspection, and discrete quality measurement. QIF supports defining or importing the product definition and reusing data for inspection planning, execution, analysis, and reporting.

QIF uses terminology and semantics from the inspection world to represent the various elements in the QIF specification. The QIF information models are normalized in XML Schema Definitions (XSD). The QIF XSDs are organized into six application areas for metrology: (1) MBD, (2) Rules, (3) Resources, (4) Plans, (5) Results, (6) Statistics. The MBD (containing the product definition) is combined with measurement rules and resources definitions to generate a plan. The plan is then executed and the results are captured. Multiple results are combined to generate statistics. QIF is an information model and format that can be exported from commercial metrology applications available in the marketplace. While, QIF does not perform the task of statistics and the other metrology methods, QIF does enable the ability to put raw inspection data into a quality context that is computer-processable.

3.2 Data Traceability and Trustworthiness

3.2.1 Data Authentication, Authorization, Traceability

In the regulated U.S. aerospace industry, the Federal Aviation Administration (FAA) requires that aerospace manufacturers to define a plan and receive FAA approval for managing and maintaining electronic design data (e.g., 3D CAD models, digital parts lists) used in the certification process [54]. Then, a parts manufacturer must be able to, "[determine] the quality, eligibility, and traceability of aeronautical parts and materials intended for installation on U.S. type-certificated products and articles," to ensure compliance with applicable regulations [55]. This requires the manufacturer to know the correct type-certificated design data and if that data was used during production. Accomplishing this task is easier said than done. Today, the traceability process is often done with significant human capital and minimal-to-no automation.

The literature reviewed supports the FAA requirements. For instance, from the perspective of product design and manufacturing, traceability is defined as the ability to discover the history of decisions in the lifecycle, control the quality of data, products, and processes, and understand the relationship between assets [56, 57, 58, 59, 60]. Tracing dependency links between assets supports establishing relationships between those assets [57]. Understanding those links and relationships helps determine how decisions made during the creation and modification of assets affect related assets. Therefore, traceability may be considered a critical quality attribute intended to ensure system outputs conform to stakeholder requirements [58, 59].

Ouertani et al. [60] suggests the following questions must be answered to support data traceability:

1. What product knowledge is created or represented?

2. Who are the actors playing different roles in creating, using, or modifying product knowledge?

3. Where is the product knowledge created and located?
4. How is the product knowledge being created or modified?
5. Why was certain product knowledge created or modified?
6. When was the product knowledge created or modified?

Therefore, data traceability cannot be separated from authentication and authorization. Authentication is the act of determining that an entity (e.g., person, data) is as the entity is declared. For example, Public Key Infrastructure (X.509-PKI) is often used to guarantee a user is authentic. In contrast, authorization is the process of determining what permissions an entity is granted by a trusted source. For example, authorization methods could define how data can be used in a defined process. In manufacturing, contracts between organizations typically define what data is declared to be and how to confirm the data declarations (i.e., authentication). However, authorization requirements are not negotiated typically such that a data user could know how data should be used during a prototype versus a production run.

Ensuring complete data integration of authentication, authorization, and traceability is important to manufacturing industries. Those organizations must be able to determine data declarations, who did what to the data, when they did it, and potentially why it was done. Both regulated and non-regulated industries need effective and efficient processes for data authentication, authorization, and traceability. Regulated industries (e.g., aerospace, automotive, medical) focus significant resources on data authentication, authorization, and traceability to ensure they comply with the appropriate public-safety oversight. Manufacturers in both regulated and non-regulated industries care about data authentication, authorization, and traceability to reduce product-liability exposure within their supply chains and in the public realm.

The cost of achieving data authorization and traceability is thought to outweigh the benefits in paper-based systems [56]. As far back as 2006, reports showed major original equipment manufacturers (OEMs) were outsourcing 60 percent to 80 percent of their manufacturing [61]. Today, the majority of OEMs are manufacturing even less product in-house – relying more on their external supply chains. For example, the Boeing 787 (Dreamliner) has 30 tier-one suppliers, which in turn contract to hundreds of tier-two and tier-three suppliers [62]. Aside from the communication challenges that come with drawing-based systems, tracing what data is being used by whom and for what purpose is costly and inefficient for the Boeing 787 program. Moreover, knowing and ensuring that the data being used is the actual FAA-approved data is a real problem. This is why Boeing made the decision to switch to a MBE to define and certify the aircraft using only 3D CAD models. However, 3D CAD models still lack commercial-off-the-shelf support for authentication, authorization, and traceability. This is the motivation for the research into using embedded X.509 digital certificates in 3D CAD models for authentication, authorization, and traceability of product data.

3.2.2 Regulatory Needs

Regulated industries are those that must comply with a government mandate and/or law. Examples of regulated industries are Aerospace (regulated by the FAA), Medical (regulated by the Food and Drug Administration (FDA)), and Consumer Products (regulated by the Consumer Product Safety Commission (CPSC)). Although, non-regulated industries do not have such mandates, they often voluntarily self-regulate through the use of consensus-based standards and testing methods. Examples of standards and testing methods are those produced by American Society of Mechanical Engineers (ASME) and Underwriters Laboratory.

The traceability process is often done with significant human capital and minimal-to-no automation [63]. For example, in the regulated U.S. aerospace industry, the FAA requires that aerospace manufacturers define a plan and receive FAA approval for managing and maintaining electronic design data (e.g., 3D CAD models, digital parts lists) used in the certification process [54]. Then, a parts manufacturer must, "[determine] the quality, eligibility, and traceability of aeronautical parts and materials intended for installation on U.S. type-certificated products and articles," to ensure compliance with applicable regulations [55]. If the manufacturer cannot provide traceable evidence that the applicable requirements were followed, then the FAA has the authority to halt production operations and to ground affected aircraft. Therefore, industry must remain cognizant of the steep consequences of poor traceability.

While industry must focus on traceability and deploy processes to manage their operations in accordance to requirements, the significant human capital burden can no longer remain the generally accepted practice. Industry, whether regulated or not, needs methods and technology to supplement the human resources assigned to ensuring effective and efficient management of data trustworthiness and traceability.

3.2.3 Product Data Quality

Product-data quality (PDQ) must be a crucial focus to ensure successful data authentication, authorization, and traceability. Product data represents product-related specifications and is typically defined using a CAD system [64]. There are two uses of product data: (1) lateral direction and (2) vertical direction [64]. Lateral direction means using product data within a phase of the product lifecycle. Vertical direction means reusing product data in subsequent product lifecycle phases. PDQ is important to both uses.

Estimates show a significant number of engineering change orders and CAD re-modeling hours are the result of error, ambiguity, and data that is unusable by downstream applications [61, 65]. A manual healing process is used typically to reach the intended quality level. The two types of healing are repair (e.g., partial restoration for improving invalid data) and rework (e.g., disposing of and remaking the whole data set) [64]. Historically, the use of computer-aided systems to represent products promised effective and efficient communication

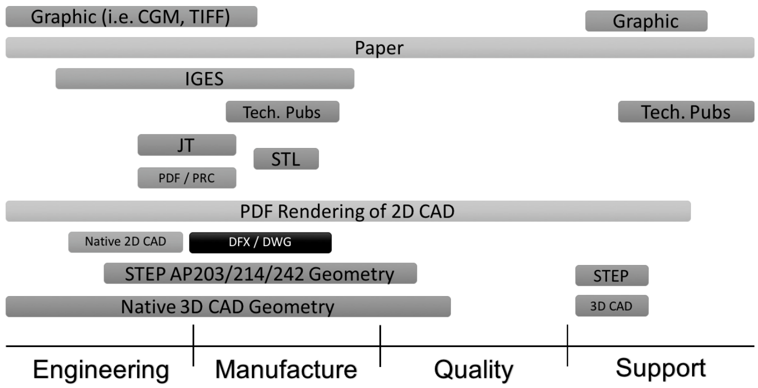


Figure 3.1: Landscape of data formats used for product-data exchange

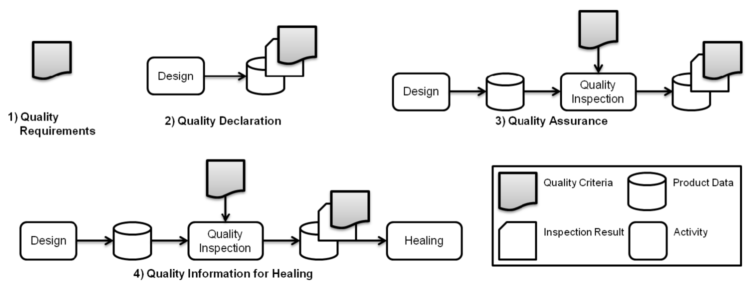


Figure 3.2: PDQ information usage scenarios (from [64])

of product data across the product lifecycle. Figure 3.1 shows the various data formats used during product-data exchange in the product lifecycle. However, promised benefits have not been achieved fully due to quality, technology, and cost limitations.

Data-interoperability formats (e.g., JT [66], PDF/PRC [45], STEP AP242 [43]) address the technology and cost factors. There is also commercial support for quality, but quality is less understood by industry than data interoperability. Again, if the translation or transfer of product data induces an error, then the product data becomes unusable for interfacing between various applications in the product lifecycle. Verifying PDQ at the point of creation ensures a known level of quality. Then, PDQ can be validated and traced throughout the product lifecycle to safeguard the data as it is exchanged, translated, inferred, and augmented. This approach enables a strong industrial verification and validation (V&V) strategy.

Kikuchi et al. [64] suggests a set of scenarios for PDQ information use (see Figure 3.2). I suggest re-purposing the Kikuchi et al. [64] scenarios as a PDQ workflow. The workflow would be four steps: (1) define PDQ requirements, (2) declare the PDQ level, (3) conduct PDQ assurance, and (4) report PDQ information.

PDQ requirements are defined PDQ criteria related to the tolerance and accuracy of product data. The PDQ requirements should also detail the V&V-diagnostic-algorithm needs. The requirements must be defined and communicated independently of the product data to ensure they are unbiased requirements. This mitigates risk of data-quality defects and economic loss from any required repair and rework. The requirements must be easily extensible to support a wide range of product types. Various standards [67, 68, 69, 70] exist that define quality requirements. Such are generally geared to an industry sector (e.g, automotive, aerospace), but are a good source for unbiased PDQ requirements.

PDQ declaration is information attached to product data that declares the PDQ level the product data satisfies. The product-data creator would declare the PDQ level. The declaration identifies what PDQ requirements are used during the V&V process. The PDQ information would also be transferred together with the product data to receiving systems. An example of PDQ level is the three technical-data-package levels (i.e., conceptual, developmental, and production) defined in MIL-STD-31000 Revision A [70]. These levels would align with the PDQ requirements defined for each level to ensure the product data is as it is declared and satisfies usage expectations.

PDQ assurance is a diagnostic test of the product data against specified PDQ requirements. The PDQ information related to quality assurance would ensure the PDQ requirements from the PDQ declaration are satisfied. The PDQ workflow step would conduct quality activities in the cyber-space similarly to the way industry conducts quality activities on physical products. The quality-assurance information would also be transferred with the product data. This supports authentication and authorization of the product data throughout the product lifecycle.

PDQ information reporting is the reporting of the PDQ results from the PDQ assurance step. If defects are discovered during diagnostic testing, the level of defect severity and any healing methods used to correct defects would also be reported. This PDQ information is used to present what quality defects were detected, the exact location of defects in the target product data, and the seriousness of the defects. The PDQ information reporting should contain information about the product data, a link to the PDQ requirements, a description of the diagnostics algorithms, and the defect information (e.g., error location, type, severity).

This workflow would run in support of the V&V strategy previously discussed. The workflow could be run during product-data creation to verify the data meets PDQ requirements. Then, the workflow could be run after exchanging or translating the product-data, to ensure the output conforms to both the input and PDQ requirements. Thus, every stage of the product lifecycle may confidently take full advantage of interfacing with the product data with traceable PDQ information. Consequently, reuse of product data significantly reduces cost, risk, and cycle time while increasing product quality [51].

3.2.4 X.509 Certificates

The X.509 standard [6], titled *Information Technology – Open Systems Interconnection – The Directory – Part 8: Public-key and Attribute Certificate Frameworks*, was first published in 1988. The Telecommunication Standardization Sector of the International Telecommunication Union (ITU-T) developed the standard, first as a recommendation, intended as the authentication framework for the X.500 series of electronic directory services. The latest version of X.509 was published in 2014 by the International Standards Organization (ISO) under standard number ISO/IEC 9594-8:2014 [6].

The X.509 standard normalized two concepts for authentication and authorization. The first is X.509-PKI [71] and Privilege Management Infrastructure (X.509-PMI) [72]. X.509-PKI addresses authentication and X.509-PMI addresses authorization.

Figure 3.3 displays the basic components of X.509-PKI (3.3a) and X.509-PMI (3.3b). The purpose of X.509-PKI is to create and manage digital certificates – primarily for authentication with a certificate authority at the top of a certificate hierarchy. The hierarchy consists of hardware, software, people, policies, and procedures [73]. Common implementations of X.509-PKI today use asymmetric (public) key encryption, where a user is issued both a private key that is only known to the user and a public key that is known to everyone [73]. X.509-PKI is the most familiar certificate infrastructure used by end-users.

X.509-PMI is less known to end-users. X.509-PMI is similar to X.509-PKI, except X.509-PMI is used for authorization. The purpose of X.509-PMI is to manage user authorizations with an attribute authority at the top of a certificate hierarchy [73]. The attribute authority references an X.509-PKI identity and delegates privileges to the identity based on the assigned privileges from a "source of authority." The attribute authority issues an "attribute certificate" that is linked to the identity provided by the X.509-PKI-based certificate. Adoption of the X.509-PMI in practice has been minimal with only a few commercially available applications.

In practice, X.509-PKI is implemented significantly more than X.509-PMI. X.509-PKI enjoys a broad range of applications – most notably Secure Sockets Layer (SSL)/Transport Layer Security (TLS) encryption of websites and Secure/Multipurpose Internet Mail Extensions (S/MIME) signing/encrypting of emails. However, X.509-PMI has seen minimal-to-no commercial adoption since its introduction to the X.509 standard in 2001. This is, in part, due to the rise of service-oriented architectures (SOAs) and attribute assertions via the Security Assertion Markup Language (SAML) specification [74] developed by Organization for the Advancement of Structured Information Standards (OASIS) [75].

X.509-PKI can be extended to include authorization information by embedding additional

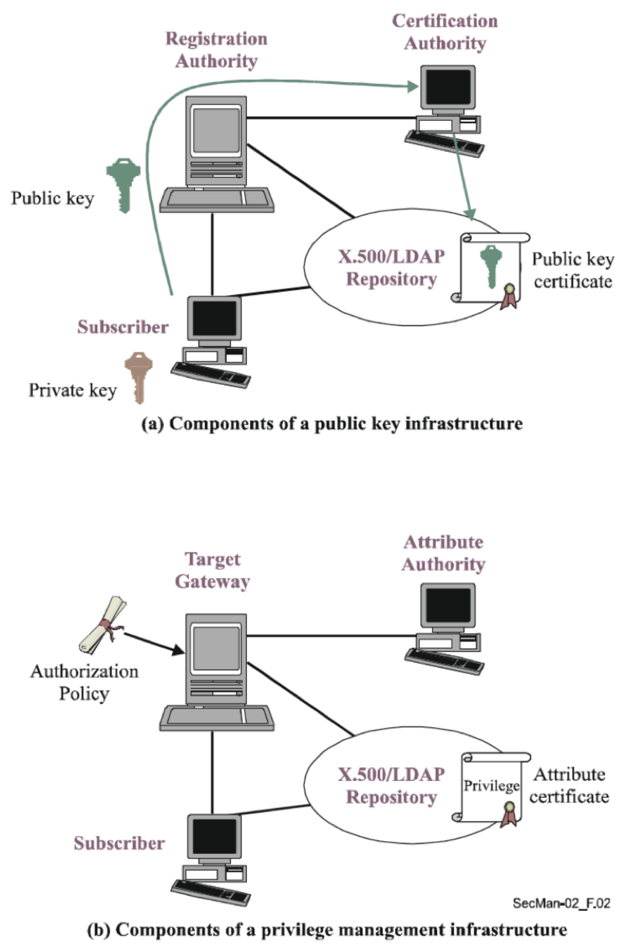


Figure 3.3: X.509 components of public key infrastructure and privilege management infrastructure (from [73])

metadata in signatures to describe privileges. The research uses X.509-PKI primarily and includes additional privilege metadata to manage authorization requirements. Taking this approach enables us to simplify the implementation of X.509 constructs while introducing traceability, authentication, and authorization to 3D CAD models.

3.2.5 Alternative Technologies

Steve Jobs said [76], "Simplicity is the ultimate sophistication." The goal was to make implementation and use for the end user as simple as possible to alleviate the need for understanding complex interactions between various actors, systems, and technologies. I investigated two alternative solutions in addition to the solution I present in this paper. The solution utilizes X.509-based digital certificates embedded in 3D CAD models for the purposes of authentication, authorization, and traceability. In addition to the solution I describe in this paper, I investigated two alternative solutions, which were:

1. Brokered data-exchange mechanisms
2. Cloud-based and software-as-a-service (SaaS) product-data repositories

Brokered data-exchange mechanisms, such as Secure File Transfer Protocol (SFTP) and Hypertext Transfer Protocol over Secure Sockets Layer (HTTPS) portals, have been a long-term solution for industry. Brokered data-exchange mechanisms are based on stable technology that has been in existence for decades. These types of mechanisms are usually continuously available and support on-demand access. They also support a simple distribution of data across the supply chain and can be centrally managed. However, brokered data-exchange mechanisms lack support for authentication, authorization, and traceability of product data unless metadata is added explicitly to the data files stored within the system. Brokered data-exchange mechanisms also require a large user, data, security, and intellectual property (IP) management overhead to ensure users can access only the data each is authorized to use. Lastly, with brokered data-exchange mechanisms, there is no control of the product data once the data is downloaded. Overall, data-exchange processes are manual and require a large management overhead, which makes brokered data-exchange mechanisms a poor choice for authentication, authorization, and traceability.

Cloud-based and SaaS product-data repositories are the newest alternative that I investigated. This alternative is implemented with the use of a proprietary application that is installed on client systems and interfaces with a centralized data repository to control the usage of product data. The cloud-based and SaaS solutions provide direct support of data authorization; they also provide continuously available data repositories and support on-demand access. These systems typically wrap product data in a proprietary-format container to keep end-users from accessing the product data without the required proprietary application being installed on client systems. Cloud-based and SaaS solutions require constant connections between client and server systems to ensure the product data is accessible. This means systems must remain “on-line” at all times, which adds a layer of unneeded complexity to interfacing with the product data. In addition, cloud-based and SaaS solutions lack standards support because the technologies are still emerging and do not have wide-spread industrial support. The lack of standards limits the industrial scalability of the cloud-based and SaaS solutions.

The shift to distributed manufacturing in retail and commercial industries is reducing traceability of product requirements. Geographically decentralizing manufacturing assets results potentially in product data being dispersed across an entire network of internal and external suppliers. In this case authentication, authorization, and traceability are imperative to ensure the right data is used at the right time. Further, the globalization and commoditization of manufacturing in the aerospace, automotive, medical, and similar industries is increasing both regulatory and data-management burdens. The best solution supports authentication, authorization, and traceability without adding to the existing burdens.

I chose to use X.509-based digital certificates to implement the strategy for authentication, authorization, and traceability because X.509 puts forth a simpler solution than the alternatives. This simpler approach supports wider opportunities for industrial adoption and scalability. In addition, the solution is enabled by the widely adopted X.509 standard.

3.3 Manufacturing-Related Concepts

Various standards and technologies exist for industry to connect and/or integrate data and systems. However, each standard and technology is often built for a specific purpose and may not apply to all viewpoints of the product lifecycle. For instance, Unified Modeling Language (UML) [77] and Systems Modeling Language (SysML) [78] are used for architecture and system modeling. CAD, CAM, computer-aided inspection (CAI) are used to generate planning and specification models in design, manufacturing, and inspection, respectively. The tools to generate each model vary widely both intra-domain and inter-domain. Standards, such as STEP [79] and Jupiter Tesselation (JT) [80], enable file-based data exchange between domains, but they have been deployed primarily in limited design contexts.

Then, each product lifecycle domain also has its own type of client support systems for managing the models built within each phase of the product lifecycle. Examples of these systems are product-data management (PDM), manufacturing execution system (MES), enterprise resource planning (ERP), and quality management system (QMS). There are multi-million dollar market sectors built around configuring, customizing, and managing these systems. Further, standards such as Open Services for Lifecycle Collaboration (OSLC) [81] and Product Life Cycle Support (PLCS) [82] or non-standard point-to-point integrations all assume the same schema and/or behavior can be used across these systems, which is impractical in today's manufacturing environments. The literature propose and commercial vendors offer centralized data repository solutions, but these types of approaches quickly breakdown under the intense burden of managing and reconciling all the data flowing in and out of the repositories. One estimate for the cost of digitally connecting and managing all artifacts in one program across its lifecycle using the tools available today is approximately $80 to $180 billion [83]. The reason for the high cost is the tools today do not support effective linked-data and require significant amounts of manual intervention to maintain. Industry needs a capability for linking all the different models and systems in distributed and universal ways to enable rapid data curation, query, discovery, and retrieval. Industry would benefit from the Semantic Web being applied to manufacturing – forming a sort of Engineering and Manufacturing Internet.

The digital thread concept shows promise for supporting industry's needs. The digital thread is an integrated information flow that connects all the phases of the product lifecycle using accepted authoritative data sources (e.g., requirements, system architecture, technical data package (TDP), 3D CAD models) [51, 84, 85], and project tasks. The product lifecycle is a complex heterogeneous system (or system-of-systems depending on how one draws the system boundaries). The aim of digital thread is to deploy an integration framework that brings all phases and systems of a product lifecycle together for making efficient and effective measurements of the lifecycle in support of data-driven methods. Specific interests relate to knowledge building, decision support, requirements management, and control. A major goal for enabling the digital thread is linking universally heterogeneous information systems and data sets across the various domains of the product lifecycle (e.g., design, manufacturing, quality) in dynamic ways without requiring one-to-one data mapping. An expected impact of achieving this goal is a significant reduction in the cost of deploying digital thread.

3.3.1 Graph Theory

The mathematical beginnings of graph theory can be traced back to Leonhard Euler (1736) [86], Arthur Cayley (1858) [87], and James Joseph Sylvester (1878) [88]. More recently, the advancements in graph theory and its application have been well covered over the past three decades [89, 90, 91, 92, 93, 94, 95, 96]. Specific to manufacturing, graph theory is applied to cutting tool / fluid performance evaluation, machining parameter optimization, material selection, machinability of materials, supply chain management [94, 97].

A graph is defined as consisting of a set of nodes (or vertices) $V = \{v_1, v_2, ...\}$ and set edges $E = \{e_1, e_2, ...\}$ where each edge e_k is associated with a pair of end nodes (v_i, v_j) [96]. Graphs are shown to be effective in modeling and analyzing structure and relationships of systems, networks, functions, and concepts [94]. Expressions exist for quickly enumerating the number of graphs that can be formed from a given set of nodes and edges. In a worst case scenario, the number of graphs that may be formed from a n number of labelled nodes is $2^{\frac{n(n-1)}{2}}$ for undirected graphs and $2^{n(n-1)}$ for directed graphs.

Table 3.1: Strengths and weaknesses of various database types and their suggested use in product lifecycle management (PLM). Adapted from [98].

Database Type	Strengths	Weaknesses	Suggested use in PLM
Relational	Known data layout and structure	Variable and hierarchical data	Transactional data in specific models
Key-value Pairs	Little or no need of indexes	Create, read, update, and delete and miscellaneous queries	Vaulting. Media.
Columnar	Horizontal scale. clustering.	Undefined data use patterns	Suppliers access. Design collaboration.
Documents	Unknown data structure	Joins and relationships	Vaulting. Media.
Graph	Flexible types of relationships	Limited scale, query-ability	Configurations. Product structure.

Assuming there is only one node per each domain of the product lifecycle[1] that can be connected, the expressions for enumerating the number of undirect and directed graphs show there could be between 1,024 (2^{10}) and 1,048,576 (2^{20}) graphs generated. While a real-world manufacturing example probably has more nodes than five, the considerable range of possibilities shown here is a significant risk for introducing uncertainity into the product lifecycle. Trying to manually manage connections of data across the product lifecycle is incomprehensible and a prime reason for the many challenges industry faces today. While graph theory applications to engineering receive sizable attention in the literature, product lifecycle management (PLM) is one area where graphs have not been significantly studied.

A reason for the lack of graph-based research in PLM is because the majority of research is still focused on data management in manufacturing [97]. However, interest in bringing "smart" technologies to manufacturing is motivating studies in graph theory applied to PLM viewpoints. Shilovitsky [98] bridged the gap between data management and PLM by suggesting different types of database technologies for use in PLM. Table 3.1 presents the strengths, weaknesses, and suggested PLM uses for five types of databases. Graphs are suggested for dealing with configurations and product structure, which aligns well with the types of relationships that must be managed as data is shared throughout the product lifecycle. The work presented in this paper accepts Shilovitsky's suggestion for using graphs in PLM to propose a method for connecting, discovering, and retrieving data across the product lifecycle.

Contextualizing data from across the product lifecycle to make design decisions is challenging because data use varies based on the role that is interacting with the data [99, 100]. Graphs can overcome some of the challenges by managing different contextual viewpoints based on what role is using the data. Sub-graphs can be extracted from the graph to enable observing the connections that matter most to a role. Further, trees – connected simple graphs where a walk starts and ends at the same node – can be extracted from graphs to enable decision making and sorting [96]. The root of the trees changes dynamically based on the domain

[1] Defined here as (1) marketing, (2) engineering, (3) manufacturing, (4) quality, and (5) sustainment.

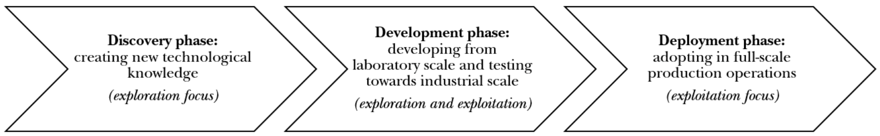


Figure 3.4: Three-phase process definition for technological innovation (based on [105])

expert's required context and the types of decisions he/she would need to make. Fortunately, various sort, search, reduction, and decision algorithms for "decision trees" and other types of graphs are widely available to solve large, computationally intensive, practical problems that are often encounter in engineering contexts [96].

3.3.2 Technological Innovation

Knight [101] proposed technological innovation means an organization has adopted a new concept beyond the generation stage of the concept. Porter [102] suggested technological innovation is a "new way of doing things that is commercialized." Freeman and Soete [103] said, "an innovation in the economic sense is accomplished only with the first commercial transaction involving new product, process system, or device..." Tidd and Bessant [104] agreed innovation is the process of growing inventions into practical use. A diagram of the technological-innovation process based on Hollen [105] is shown in Figure 3.4. The literature [105, 106], both recent and past, show technological innovation as a three-step process of discovery, development, and deployment.

The first phase in the technological-innovation process is discovery. This phase may be considered synonymous with invention. New knowledge is created during the discovery phase. The output from the discovery phase is typically a conceptual design from a Research and Development (R&D) activity.

The second technological-innovation phase is development. This phase is a transition activity. In product development, the conceptual-design task is transitioning towards detailed-design activities. Management of technological innovation is important during the development phase because successful commercialization depends on the maturity level of the technology. The output of the development phase is a complete definition for the technology.

The third phase is deployment. This phase is where a process is being deployed to production operations, or products are available for delivery to the marketplace. Development is complete or near completion when the deployment phase begins. The output of the last phase is a new and complete technology.

Management must remain a critical focus during the deployment phase because many scholars consider the commercialization of technology the least managed activity in the technological-innovation process [107]. The methods used to commercialize and market technology significantly influence the success or failure of products [108]. Products with newly-commercialized technology fail at a rate of 40 percent to 50 percent [107]. The demonstrated importance of management and decision-making is the motivation behind this paper.

3.3.3 Managing Decisions for Creativity, Development, and Change

"Creativity is the production of novel and useful ideas in any domain" [109]. Amabile [109] proposed a model of creativity that requires abilities in three major components, which are expertise, creative thinking, and intrinsic task motivation. The combined skills in each category enable creativity. The field of psychology teaches that anyone is capable of creativity, but the level of creativity is enhanced or limited by interactions with the social environment.

Lewin's Equation [110], $B = f(P, E)$, proposed behavior (B) is a function of interactions between people (P) and their environment (E). Following this idea, I argue innovation is a function of a person's creative ability and his/her interaction with the social environment. Further, Hoegl and Parboteeah [111] suggested that the quality of team collaboration influences the utilization of the teams' technical skills and directs those skills toward the critical-performance dimensions.

Considering, Hoegl and Parboteeah [111], I propose extending Lewin's Equation [110] to organizations by arguing that innovation is a function of the organization's overall creative ability and its social interactions within the environment. That is $I = f(\sum P_i, E \in O)$, where I represents innovation, i represents individuals in the organization, and O represents the organization. Therefore, managing and encouraging creativity at the personal level should support a positive environment for innovation at the organizational level.

Amabile [109] argued that individuals with basic capacities can develop moderately creative solutions to some problems some of the time. However, challenging problems of high importance require subject matter experts with extensive knowledge in the field of work. A baseline level of expertise in the engineering domain is needed to ensure the ideas produced by the creative process are "novel and useful" [109].

Amabile's [109] and Hoegl's and Parboteeah's [111] conclusions support Cooper's [112] recommendations for including all critical roles in a product-development process from the start of the process. Cooper further suggests there are two ways to succeed in innovation – (1) doing projects right and (2) doing the right projects. Doing projects right requires a process to follow commonly accepted management guides. These guides should include using teams effectively, doing up-front research before starting development, analyzing the voice of the customer, and ensuring a stable product definition prior to deployment or launch. Doing the right projects requires the "right" expertise to know what the right portfolio of projects looks like. This relates to Amabile's [109] conclusion that a basic level of expertise is needed to determine if something is "novel and useful."

Cooper also developed a stage-gate process model that breaks the product-innovation process

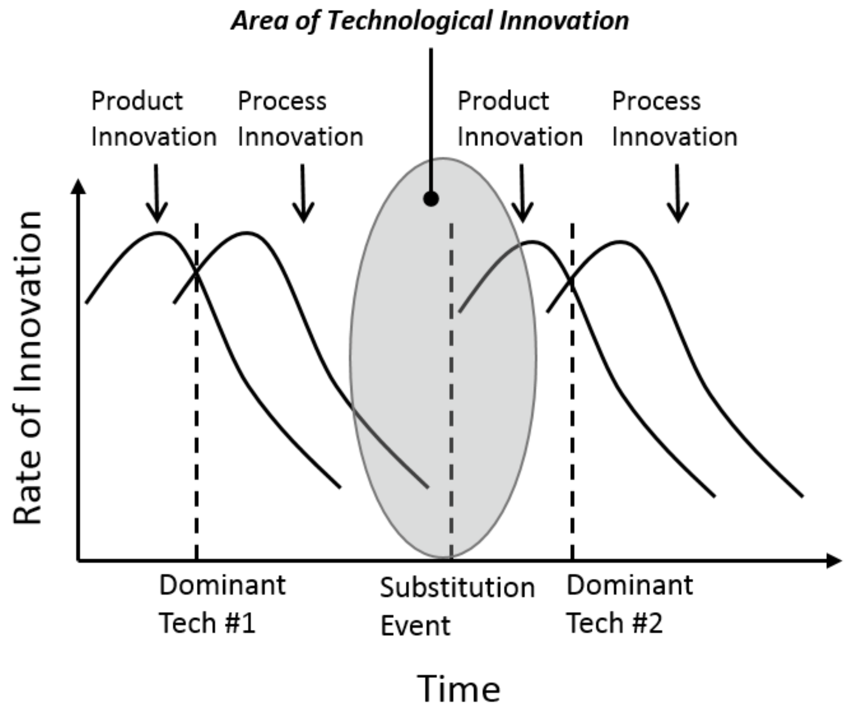


Figure 3.5: Technology cycles and technological-innovation (based on [113])

into five stages, each requiring the passage of a gate before proceeding to the next stage. The gates provide quality control to the process by incorporating go/no-go decisions at strategic points in the process. While Cooper's model provides a good foundation for managing product-development activities, it may fall susceptible to disruptive changes that could occur during the activities – specifically changes due to the technological-innovation process. This opens up Cooper's model to the risk of pursuing decisions that are no longer the right decisions.

Manufacturing organizations operate in an environment of constant change. Organizations must be prepared to manage the changes through effective decision-making. Managing changes effectively is an important part of ensuring sustainable success within an organization. Organizational strategies, structures, skills, and cultures must evolve over time to reflect changes in markets and technology [113].

Specifically related to technology, change happens in cycles [113]. These cycles are best explained with an illustration presented in Figure 3.5. Technology cycles begin with high rates of innovation until a dominant technology emerges. As technology matures, the rates of innovation slow. As competition continues in the market, eventually new technology needs to be developed to sustain success. This forces a rapid increase in the rate of innovation – leading to substitute technologies via the technological-innovation process.

In the manufacturing domain, data are being used in new ways that are beginning to enable near-real-time decision making. Data-driven decision making is at the core of Industry 4.0, Industrial Internet of Things (IIoT), and Smart Manufacturing strategies. Significant innovations in data-driven techniques were achieved in the 1980s, but other technological innovations were dominant at that time. As the 2000s approached, the rate of manufacturing-related technical innovation decreased, causing manufacturing to look for new avenues to grow and increase productivity. Manufacturing is again in a time of increased innovation and the shaded area of technological innovation shown in Figure 3.5 is imminent. New technologies and new integrations of technologies are revolutionizing the way manufacturing is conducted.

3.3.4 Control Theory Related to Manufacturing Decisions

Control means measuring a quantity or condition in a system and applying a determined quantity or condition to the system to correct or limit the deviation of the measured value from a desired value [114]. Using the word "system" in control problems refers typically to a representation of the actual thing that someone is trying to control.

In engineering, mathematical modeling is a common way of representing a system for controls analysis [114]. Modern control theory has become popular for analyzing complex systems, which often have multiple inputs and outputs as parts of the overall system [114]. A popular method for analyzing these types of complex systems is state-space analysis [115].

In this work, without pretension of being exhaustive, I was less interested in the formulation of representative models. My interest is in developing a foundational structure to describe the behavior of the system completely at any point in time. That is important for being able to accurately assess the decision-making process. This is why there is an interest in control theory – specifically state-space analysis.

While modern control theories provide great values to the engineering domain, they tend to lack complete diagnostics to facilitate controlling the decision-making process. Control must also be reviewed in the contexts of management and human-factors. Management-control systems include human-resource tools. Organizations might employ management-control techniques in budgets, rules, operating procedures, and performance-appraisal systems to help gain control over employee behaviors [116].

Performance-appraisal systems may include goal setting, which is important to achieving organizational objectives [117]. Organizations implement goal setting with employees because studies show goal setting supports positive motivation and contributes to improved employee performance [118]. Goal setting has also been shown to create competition amongst employees and teams, which increases motivation throughout an organization [116] and improves decision-making processes [118].

Since the 1960's, organizations have used Drucker's [119] work, "Management by Objectives," to control behaviors. Drucker's work has five steps: (1) define organizational objectives, (2) set worker objectives, (3) monitor progress, (4) evaluate performance, and (5) reward results. In the first step, management describes the organization's vision and objectives to the employees. In the second step, each employee meets with management to set specific goals for the employee. The third and fourth steps relate to monitoring and measuring the progress of each employee's goals and providing an evaluation at the end of the performance period. In the last step, the organization rewards each employee based on his/her results.

In Drucker's theory, goal setting is an integral part in all levels of an organization. Ceresia [118] suggested robust management control is supported by both taking into account Drucker's guidance and ensuring positive employee motivation. However, Drucker's theory and Ceresia's recommendations also lack guidance in continuously assessing organization objectives and goals.

Simon [120] published directly on the topic of using control systems to drive strategic renewal. He defines management control systems as "formal, information-based routines and procedures managers use to maintain or alter patterns in organizational activities." Simon also outlines a business strategy with four variables that require assessment. He called these variables "levers of control," which he defined as belief systems, boundary systems, diagnostic-control systems, and interactive-control systems.

I am most interested in the diagnostic-control-systems lever, which provides controls in an optimal spot of the organization because input controls and process standardization do not provide diagnostic management. Input controls maximize creativity but increase risks to cost controls, while organizational goals and standardization minimize creativity and innovation. Diagnostic control systems monitor organizational outcomes, which get compared against important performance dimensions of a strategy. Simon called these "critical performance variables."

In manufacturing industries, critical performance variables are called key performance indicators (KPIs). Simon suggested using KPIs to track the probability of meeting goals or the largest potential for gain over time. These categories of KPIs are considered effectiveness criterion and efficiency criterion, respectively.

The standard ANSI/ISA 95 [121] provides guidance to integrating control systems into enterprise hierarchies. The standard describes a pyramid hierarchy starting with an enterprise level at the top, then moving down to an operations-management level, then a sensing and control level, and finally a devices level. The standard, itself, focuses on the operations-management level.

In Figure 3.6 I combine the work of Ogata [114], Drucker [119], Simons [120], and ANSI/ISA 95 [121] to form a model for strategy diagnosis in a manufacturing-enterprise-control-system integration. This model demonstrates how organizational strategies, structures, skills, and cultures could evolve according to Tushman [113]. The model depicted in Figure 3.6 provides a good foundation for controlling the strategies of organizations implementing smart manu-

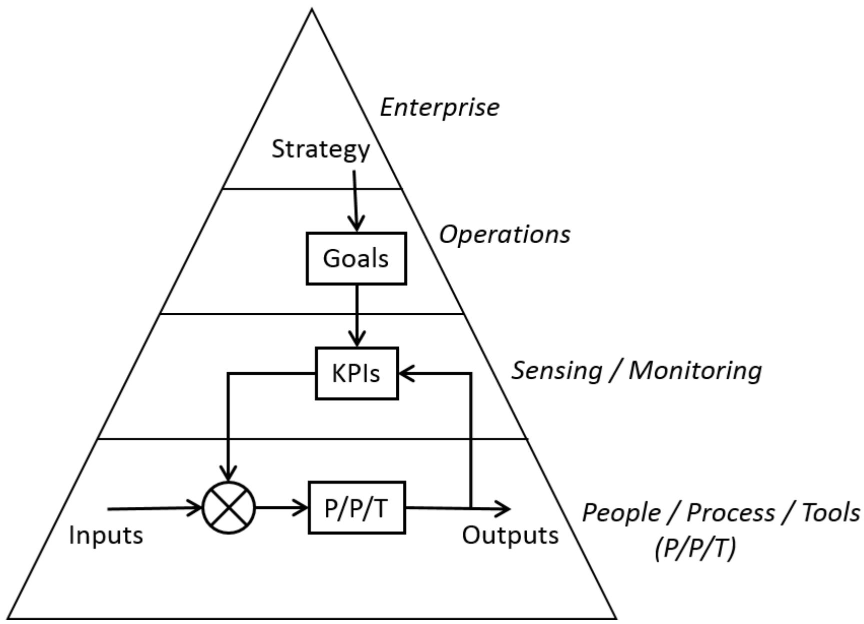


Figure 3.6: Strategy diagnosis in an enterprise-control system integration (based on [114, 120, 121])

facturing, but, like Cooper's [112] model and Drucker's [119] theory, the model for strategy diagnosis may be susceptible to the various types of change – resulting in organizations pursuing strategies that are no longer ideal.

Argyris' [122] developed the concept of double-looping learning. The concept can be represented as a control system. Examples of single-loop-learning and double-loop-learning as control systems are shown in Figure 3.7. In the double-loop example, there are two "sensors." The first sensor measures the system output in context to the local goal. The second sensor measures the system output in context to the overall goal.

In double-loop learning, the system inputs are modified based on the system output compared to the local goal, but the local goal may also be modified in light of the system output not trending toward the overall goal. The system could also be controlled by modifying the overall goal instead of the local goal.

3.3.5 Managing Lifecycles

Maturing a new product idea to commercialization requires nurturing and oversight, which only proper management controls can provide [124]. Simons [125] defines management-control systems as "formal, information-based routines and procedures managers use to maintain or alter patterns in organizational activities." Industry has applied various management techniques to all aspects of product lifecycle activities. This has ranged from project manage-

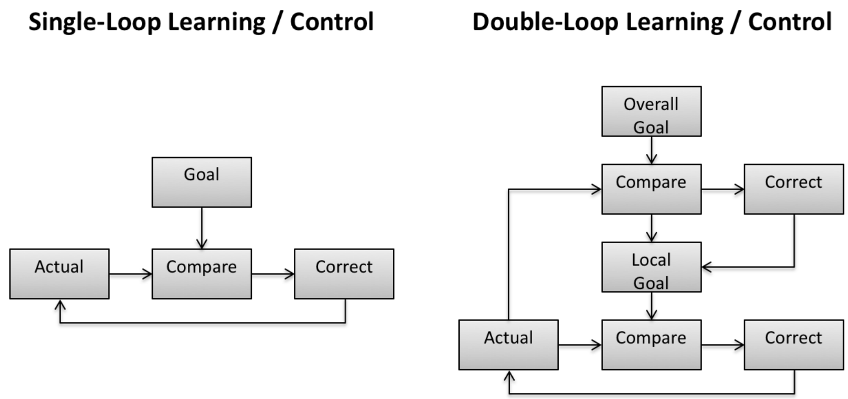


Figure 3.7: The single-loop learning process compared to the double-loop-learning process (based on [123])

ment methods (e.g., waterfall, phase gate, agile) to process management (e.g., MES, ERP) to data management (e.g., PDM). Spreadsheets, sticky notes, file shares, and whiteboards are some of the most basic management tools used by industry [126].

However, in more sophisticated approaches, industry uses the theory of PLM to manage their hardware products from the earliest ideas to the end of the products' lives [127]. PLM gained prominence in the automotive and aerospace sectors in the late-1990s and early-2000s. A goal of PLM is to enable cross-functional and cross-enterprise collaboration between participants that are geographically dispersed [127]. Unfortunately today, PLM is often conflated with the software platforms marketed by engineering-tool providers and focused primarily on the needs of engineering, which reinforces the information-silo paradox that limits knowledge exchange across the entire product lifecycle [128]. Conversely, industry requires management concepts and infrastructure for integrating heterogeneous systems that scale across several domains to support leveraging the best available data possible [99, 129].

Whereas PLM focuses primarily on hardware products, application lifecycle management (ALM) focuses on software. The intent behind ALM is to manage software development from requirements through coding to testing, release, and maintenance [126]. Agile development methods have increased the importance of ALM [130]. Similar to PLM, tools dedicated to ALM tend to be focused on a single discipline or function [126]. While focused on software, ALM suffers from many of the same challenges facing PLM.

Converging and/or integrating PLM and ALM is gaining attention. Both management techniques can be grouped into three categories: (1) governance, (2) development, and (3) operations [131]. Governance is the overall policy and management of the lifecycle as it relates to the commercialized product. Development is the management of the product development cycle. Operations is the using, monitoring, and maintaining of the product. Given the rise of "smart" products and processes, coupled with industrial internet of things (IIoT), combining

PLM and ALM into a unified method is the next logical progression of lifecycle management because smart products require mechatronic specifications (i.e., designs composed of mechanical designs, electrical designs, and embedded software). In managing these types of products, engineers need harmonized workflows, integrated systems, and redundancy-free data reuse [132]. These needs give credence to using service-oriented architectures in PLM.

In using service-oriented PLM systems, three key requirements must be satisfied by those systems [133]:

- The system must support interactive query, discovery, and retrieval of information from several different sources across the product lifecycle
- The system must provide near-real-time information and/or notify the user if information is out-of-date or when an update is available
- The system must integrate the user's context and expand the functionality of existing tools instead of providing a completely seperate tool

Further, service-oriented architectures provide a significant integration benefit over point-to-point integration. Point-to-point integration of tools is fragile and expensive to develop and maintain because a ripple affect of changes occurs as one tool is modified or replaced [126]. Context is also harder to manage during point-to-point integrations because the individual tools are centered on one discipline while the integrations must support multiple viewpoints, which could lead to the deployment of multiple point-to-point integrations for connecting a single tool to a suite of other tools. Conversely, service-oriented architectures offer the benefit of composing systems dynamically to meet changing demands in the operation of manufacturing systems [134].

3.3.6 System and Information Modeling

The Core Product Model (CPM), and later the Revised Core Product Model (CPM2), was developed by the National Institute of Standards and Technology (NIST) with the goal of providing artifact representation for product-development information that encompasses a broad range of engineering design concepts including geometry, function, form, behavior, material, physical and functional decomposition, mappings between function and form, and various kinds of relationships [135, 136]. The argument for CPM was there is a need for, "formal representation, capture, and exchange of the entire range of information generated and used in the product development process, not just of the representation of the product resulting from the completion of the design process," [135]. The state-of-the-art for information exchange at the time of CPM2 was direct electronic interchange (e.g., the transmission of files between organizations) and concepts like linked data [137] did not exist yet. A problem with CPM2 is it takes a near-monolithic model approach whereas all phases and functions of the product lifecycle must use the same schema. CPM2 saw little adoption because it was quickly overshadowed by linked data concepts coupled with the semantic web [138, 139].

However, CPM did a good job in documenting the information requirements across the product lifecycle. While the recommendation to use an all-encompassing model did not work out, the introduction of link data increased interest in applying ontological analysis to the the information requirements from CPM. It is understood that different roles across the product lifecycle have different context and viewpoints [99, 100]. Semantic representation of product-related information, using ontologies, would enable multiple viewpoints to leverage data across the product lifecycle [140]. Lee et al. [140] proposed an ontological framework for semantic product modeling. The proposal outlined three requirements for a generic product model to include in a product ontology:

- The generic product model must be readily specialized for specific products
- The generic product model must provide information to all stakeholders in the product lifecycle (e.g, designer, manufacturer, maintenance technician)
- The generic product model must provide explicit, logical semantics of the concepts and relationships to the stakeholder without requiring ontology expertise

Lee et al.'s proposed a multilevel information framework that includes a level for metamodels, product models, and instance models. The semantic-based product metamodel (SPMM), built upon CPM2, was proposed for the metamodel level of Lee et al.'s framework [140]. While the idea of using ontologies and mutlilevel information frameworks is a step towards satisfying Catic's and Andersson's [133] requirements listed in Section 3.3.5, there is still the assumption that every piece of data would live in the same place or that every tool / system has or uses the same schema. This is not practical in today's manufacturing environments. Tools used by industry can not be forced to – nor can it be enforced that every tool must – use the same schema and/or behave the same way. Ontologies and schemas should be generated to solve specific problems. It is too hard to generalize everything to all types of problems. Domain-specific, purpose-built approaches enable expert systems that can effectively solve problems. Whereas, approaches using broad concepts are often not useful without adding specifics for a focal problem.

While CPM2 and the Lee et al. framework have weaknesses that make those proposals impractical for industry, they are a step toward providing significant impact to industry. The two proposals could be extended to utilize holistic system models, services, application programming interfaces (APIs), ontologies, and schemas to dynamically "link" things together across the product lifecycle. The data could live where industry wants it to live in the format that industry wants the data to be in. Then, industry could use methods that generate data observatories for interacting with the data, discovering information, and extracting knowledge.

The Total System Model (TSM) concept fills the gaps identified in this section. The TSM is a conceptual system architecture model, built upon SysML, that connects domain-specific models from across the product lifecycle while tracking the relationships of the connections [141, 142]. The goal of the TSM is to manage models, including their various versions, and the inter-model connections across different repositories such that each discipline in the product lifecycle can use purpose-built tools while utilizing product lifecycle data across federated services. Two key lessons were learned during the development of the TSM [141]. The first is simply creating connections between artifacts is not enough – effective and efficient query, visualization, and ability to apply information across the connections are required too. The second lesson is the user of the TSM needs the ability to traverse both inter- and intra-model connections without switching between different tools – thus, the results of addressing the first lesson must be integrated into the user's standard tool of choice.

However, the TSM does have two remaining gaps in its approach. The first gap is the TSM is missing a proposed way for universal addressing of artifacts. Without such a way for universal addressing, managing changes in system configurations, locations, and migration of artifacts becomes burdensome because each way of addressing (e.g., artifact pointers, protocols, service ports) must be tracked by each system participating in the TSM. The second gap is the TSM has no way of understanding what the connected artifact represents without having access to the system where the artifact resides. This assumes that each user of the TSM has the correct authentication and authorization to access those systems participating in the TSM.

I propose combining CPM2 and TSM, closing the identified gaps of the two concepts, and satisfying the requirements listed in Section 3.3.5. In Chapter 5, I will describe the approach to a global identifier (GID) for universal addressing, an identifying system similar in concept to Domain Name System (DNS), and an idea for distributed metadata repositories that enable query, discovery, and retrieval similar to popular internet search engines.

Chapter Bibliography

[1] Open Knowledge Foundation. About, 2015. URL https://okfn.org/about/.

[2] Open Knowledge Foundation. CKAN, 2015. URL https://okfn.org/about/our-impact/ckan/.

[3] CKAN Association. About, 2015. URL http://ckan.org/about/.

[4] Open Data Institute. About the ODI, 2015. URL http://theodi.org/about-us.

[5] Open Data Institute. Open data certificate, 2015. URL https://certificates.theodi.org/.

[6] Telecommunication Standardization Sector of ITU. Information technology – open systems interconnection – the directory – part 8: Public-key and attribute certifi-

cate frameworks, 2014. URL http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=64854.

[7] Open Data Institute. About open data certificates, 2015. URL https://certificates.theodi.org/about.

[8] U.S. General Services Administration. Data.gov, 2015. URL http://www.data.gov/.

[9] U.S. General Services Administration. Impact, 2015. URL http://www.data.gov/impact/.

[10] National Aeronautics and Space Administration. Physical science information: History and purpose, 2014. URL http://psi.nasa.gov/History.aspx.

[11] Consortium for Ocean Leadership. Final network design. Report, Ocean Observatories Initiative, 2010. URL oceanleadership.org/wp-acontent/uploads/2009/04/1101-00000_FND_OOI_2010-04-22_ver_2-06_public1.pdf.

[12] EarthCube. Earthcube strategic science plan. Report, National Science Foundation, 2015. URL http://earthcube.org/document/earthcube-strategic-science-plan.

[13] US Law. Department of Defense Authorization Act, 1985, 1984. URL http://www.gpo.gov/fdsys/granule/STATUTE-98/STATUTE-98-Pg2492/content-detail.html.

[14] Logistics Information Service. Military engineering data asset locator system (MEDALS). Report, Defense Logistics Agency, 2014. URL https://www.dlis.dla.mil/medals/pdfs/MEDALSHomepageOverview.pdf.

[15] Mark R. Cutkosky, Jay M. Tenenbaum, and Jay Glicksman. Madefast: collaborative engineering over the internet. *Communications of the ACM*, 39(9):78–87, 1996. ISSN 0001-0782. doi: 10.1145/234215.234474. URL http://dl.acm.org/citation.cfm?id=234474.

[16] Angus Stevenson and Maurice Waite. *Concise Oxford English dictionary*. Oxford University Press, Oxford ; New York, 12th edition, 2011. ISBN 9780199601080.

[17] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific American*, May, 2001. URL http://www.scientificamerican.com/article/the-semantic-web/.

[18] Jacco van Ossenbruggen, Lynda Hardman, and Lloyd Rutledge. Hypermedia and the semantic web: A research agenda. *Journal of Digital Information*, 3(1), 2002. ISSN 1368-7506. URL https://journals.tdl.org/jodi/index.php/jodi/article/view/78/77.

[19] Tim Berners-Lee. Linked data, 2009 2006. URL http://www.w3.org/DesignIssues/LinkedData.html.

[20] World Wide Web Consortium. Semantic web, 2006. URL https://www.w3.org/standards/semanticweb/.

[21] World Wide Web Consortium. Resource description framework (RDF), 2014. URL http://www.w3.org/RDF/.

[22] World Wide Web Consortium. SPARQL 1.1 overview, 2013. URL http://www.w3.org/TR/sparql11-overview/.

[23] Nitesh Khilwani, Jennifer A. Harding, and Alok K. Choudhary. Semantic web in manufacturing. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 223(7):905–925, 2009. doi: 10.1243/09544054JEM1399. URL https://dspace.lboro.ac.uk/2134/9519.

[24] Rui P. Fernandes, Ian R. Grosse, Sundar Krishnamurty, Paul Witherell, and Jack C. Wileden. Semantic methods supporting engineering design innovation. *Advanced Engineering Informatics*, 25(2):185–192, 2011. ISSN 14740346. doi: 10.1016/j.aei.2010.08.001. URL http://dx.doi.org/10.1016/j.aei.2010.08.001.

[25] Farhad Ameri and Lalit Patil. Digital manufacturing market: A semantic web-based framework for agile supply chain deployment. *Journal of Intelligent Manufacturing*, 23(5):1817–1832, 2012. ISSN 09565515. doi: 10.1007/s10845-010-0495-z. URL http://dx.doi.org/10.1007/s10845-010-0495-z.

[26] Farhad Ameri and Christian McArthur. Semantic rule modelling for intelligent supplier discovery. *International Journal of Computer Integrated Manufacturing*, 27(6):570–590, 2014. ISSN 0951192X. doi: 10.1080/0951192X.2013.834467. URL http://dx.doi.org/10.1080/0951192X.2013.834467.

[27] World Wide Web Consortium. Inference, 2008. URL http://www.w3.org/standards/semanticweb/inference.

[28] World Wide Web Consortium. OWL ontology web language overview, 2004. URL http://www.w3.org/TR/owl-features/.

[29] World Wide Web Consortium. SKOS simple knowledge organization system reference, 2009. URL http://www.w3.org/TR/skos-reference/.

[30] World Wide Web Consortium. Vertical applications, 2008. URL http://www.w3.org/standards/semanticweb/applications.html.

[31] Alan Ruttenberg, Jonathan A. Rees, Matthias Samwald, and M. Scott Marshall. Life sciences on the semantic web: the neurocommons and beyond. *Briefings in Bioinformatics*, 10(2):193–204, 2009. doi: 10.1093/bib/bbp004. URL http://bib.oxfordjournals.org/content/10/2/193.abstract.

[32] Kei-Hoi Cheung, H. Robert Frost, M. Scott Marshall, Eric Prud'hommeaux, Matthias Samwald, Jun Zhao, and Adrian Paschke. A journey to semantic web query federation in the life sciences. *BMC Bioinformatics*, 10(Suppl 10):S10–S10, 2009. ISSN 1471-2105. doi: 10.1186/1471-2105-10-S10-S10. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2755818/.

[33] Helena F. Deus, Eric Prud'hommeaux, Michael Miller, Jun Zhao, James Malone, Tomasz Adamusiak, Jim McCusker, Sudeshna Das, Philippe Rocca Serra, Ronan Fox, and M. Scott Marshall. Translating standards into practice – one semantic web api for gene expression. *Journal of Biomedical Informatics*, 45(4):782–794, 2012. doi: 10.1016/j.jbi.2012.03.002. URL http://dx.doi.org/10.1016/j.jbi.2012.03.002.

[34] M. Scott Marshall, Richard Boyce, Helena F. Deus, Jun Zhao, Egon L. Willighagen, Matthias Samwald, Elgar Pichler, Janos Hajagos, Eric Prud'hommeaux, and Susie Stephens. Emerging practices for mapping and linking life sciences data using rdf: A case series. *Web Semantics: Science, Services and Agents on the World Wide Web*, 14 (0):2–13, 2012. ISSN 1570-8268. doi: http://dx.doi.org/10.1016/j.websem.2012.02.003. URL http://www.sciencedirect.com/science/article/pii/S1570826812000376.

[35] Paul Witherell, Sundar Krishnamurty, Ian R. Grosse, and Jack C. Wileden. Improved knowledge management through first-order logic in engineering design ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, 24(2):245–257, 2010. ISSN 08900604. doi: 10.1017/S0890060409990096. URL http://dx.doi.org/10.1017/S0890060409990096.

[36] Dazhong Wu, J. Lane Thames, David W. Rosen, and Dirk Schaefer. Enhancing the product realization process with cloud-based design and manufacturing systems. *Journal of Computing and Information Science in Engineering*, 13(4), 2013. ISSN 15309827. doi: 10.1115/1.4025257. URL http://dx.doi.org/10.1115/1.4025257.

[37] Dazhong Wu, David W. Rosen, Lihui Wang, and Dirk Schaefer. Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. *Computer-Aided Design*, 59(0):1–14, 2015. ISSN 0010-4485. doi: 10.1016/j.cad.2014.07.006. URL http://www.sciencedirect.com/science/article/pii/S0010448514001560.

[38] Xun Xu. From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, 28(1):75–86, 2012. ISSN 0736-5845. doi: 10.1016/j.rcim.2011.07.002. URL http://www.sciencedirect.com/science/article/pii/S0736584511000949.

[39] Energetics Inc. Measurement science roadmap for prognostics and health management for smart manufacturing system. Report, National Institute of Standards and Technology, 2015.

[40] R. Gao, L. Wang, R. Teti, D. Dornfeld, S. Kumara, M. Mori, and M. Helu. Cloud-enabled prognosis for manufacturing. *CIRP Annals - Manufacturing Technology*, 64 (2):749–772, 2015. ISSN 0007-8506. doi: 10.1016/j.cirp.2015.05.011. URL http://www.sciencedirect.com/science/article/pii/S000785061500150X.

[41] Moneer Helu and Thomas Hedberg Jr. Enabling smart manufacturing research and development using a product lifecycle test bed. *Procedia Manufacturing*, 1:86–97, 2015. ISSN 2351-9789. doi: http://dx.doi.org/10.1016/j.promfg.2015.09.066.

[42] Min Li, Shuming Gao, and Charlie C. Wang. Real-time collaborative design with heterogeneous CAD systems based on neutral modeling commands. *Journal of Computing and Information Science in Engineering*, 7(2):113–125, 2006. ISSN 1530-9827. doi: 10.1115/1.2720880. URL http://dx.doi.org/10.1115/1.2720880.

[43] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014.

[44] International Standards Organization. Document management – portable document format – part 1: PDF 1.7, 2008.

[45] International Standards Organization. Document management – 3D use of product representation compact (PRC) format – part 1: PRC 10001, 2014.

[46] International Standards Organization. Automation systems and integration – numerical control of machines – program format and definitions of address words – part 1: Data format for positioning, line motion and contouring control systems, 2009.

[47] MTConnect Institute. Mtconnect standard, 2014. URL http://www.mtconnect.org/media/39542/mtc_part_1_overview_v1.3.pdf.

[48] Dimensional Metrology Standards Consortium. Part 1: Overview and fundamental principles in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014. URL http://www.qifstandards.org.

[49] Allison Barnard Feeney, Simon P. Frechette, and Vijay Srinivasan. A portrait of an ISO STEP tolerancing standard as an enabler of smart manufacturing systems. *Journal of Computing and Information Science in Engineering*, 15(2):021001–021001, 2015. ISSN 1530-9827. doi: 10.1115/1.4029050. URL http://dx.doi.org/10.1115/1.4029050.

[50] Asa Trainer, Thomas Hedberg Jr, Allison Barnard Feeney, Kevin Fischer, and Phil Rosche. Gaps analysis of integrating product design, manufacturing, and quality data in the supply chain using model-based definition, 2016. 2016 ASME Manufacturing Science and Engineering Conference.

[51] Thomas Hedberg Jr, Joshua Lubell, Lyle Fischer, Larry Maggiano, and Allison Barnard Feeney. Testing the digital thread in support of model-based manufacturing and inspection. *Journal of Computing and Information Science in Engineering*, 16(2):021001, mar 2016. ISSN 1530-9827. doi: 10.1115/1.4032697.

[52] Kevin Ficher, Phil Rosche, and Asa Trainer. Investigating the impact of standards-based interoperability for design to manufacturing and quality in the supply chain. Report, National Institute of Standard and Technology, 2015. URL http://www.nist.gov/manuscript-publication-search.cfm?pub_id=920033. NIST GCR 15-1009.

[53] Suk-Hwan Suh. *Theory and design of CNC systems*. Springer series in advanced manufacturing. Springer, London, 2008. ISBN 1848003358 (hbk.) 9781848003354 (hbk.) 1848003366 (ebook) 9781848003361 (ebook).

[54] David W. Hempe. Advisory Circular 21-48. Report, Federal Aviation Administration, U.S. Department of Transportation, 2010. URL http://www.faa.gov/documentLibrary/media/Advisory_Circular/AC%2021-48.pdf.

[55] John M. Allen. Advisory Circular 20-62E. Report, Federal Aviation Administration, U.S. Department of Transportation, 2010. URL http://www.faa.gov/documentLibrary/media/Advisory_Circular/AC%2020-62E.pdf.

[56] V. L. Hamilton and M. L. Beeby. Issues of traceability in integrating tools. In *IEEE Colloquium on Tools and Techniques for Maintaining Traceability During Design*, pages 4/1–4/3, 1991.

[57] B. Ramesh. Process knowledge management with traceability. *IEEE Transactions on Software Engineering*, 19(3):50–52, 2002. ISSN 0740-7459. doi: 10.1109/MS.2002.1003454.

[58] Kannan Mohan and Balasubramaniam Ramesh. Traceability-based knowledge integration in group decision and negotiation activities. *Decision Support Systems*, 43 (3):968–989, 2007. ISSN 0167-9236. doi: http://dx.doi.org/10.1016/j.dss.2005.05.026. URL http://www.sciencedirect.com/science/article/pii/S0167923605000916.

[59] Kannan Mohan, Peng Xu, Lan Cao, and Balasubramaniam Ramesh. Improving change management in software development: Integrating traceability and software configuration management. *Decision Support Systems*, 45(4):922–936, 2008. ISSN 0167-9236. doi: http://dx.doi.org/10.1016/j.dss.2008.03.003. URL http://www.sciencedirect.com/science/article/pii/S0167923608000523.

[60] M. Z. Ouertani, S. Baïna, L. Gzara, and G. Morel. Traceability and management of dispersed product knowledge during design and manufacturing. *Computer-Aided Design*, 43(5):546–562, 2011. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/j.cad.2010.03.006. URL http://www.sciencedirect.com/science/article/pii/S0010448510000618.

[61] J. Yang, S. Han, H. Kang, and J. Kim. Product data quality assurance for e-manufacturing in the automotive industry. *International Journal of Computer Integrated Manufacturing*, 19(2):136–147, 2006. ISSN 0951-192X. doi: 10.1080/09511920500171261. URL http://dx.doi.org/10.1080/09511920500171261.

[62] Mike Collins. The Boeing supply chain model. *Manufacturing.net*, 2010. URL http://www.manufacturing.net/news/2010/07/boeing-supply-chain-model.

[63] Thomas D. Hedberg Jr, Sylvere Krima, and Jaime A. Camelio. Embedding x.509 digital certificates in three-dimensional models for authentication, authorization, and traceability of product data. *Journal of Computing and Information Science in Engineering*, 17(1):011008–011008–11, 2016. ISSN 1530-9827. doi: 10.1115/1.4034131.

[64] Yoshihito Kikuchi, Hiroyuki Hiraoka, Akihiko Otaka, Fumiki Tanaka, Kazuya G. Kobayashi, and Atsuto Soma. PDQ (product data quality): Representation of data quality for product data and specifically for shape data. *Journal of Computing and Information Science in Engineering*, 10(2):021003–021003, 2010. ISSN 1530-9827. doi: 10.1115/1.3402615. URL http://dx.doi.org/10.1115/1.3402615.

[65] Dan Walker. Introduction to TOPGUN XI. In *2001 COE Conference*, 2001.

[66] International Standards Organization. Industrial automation systems and integration – JT file format specification for 3D visualization, 2012.

[67] Automotive Industry Action Group. Defining product data quality, 1999.

[68] Automotive Industry Action Group. Product data quality: Guidelines for the global automotive industry, 2001.

[69] International Standards Organization. SASIG product data quality guidelines for the global automotive industry, 2006.

[70] US Department of Defense. Standard practice: Technical data packages, 11/1/2009 2013.

[71] The Internet Engineering Task Force. Internet X.509 public key infrastructure certificate and certificate revocation list (CRL) profile, 2013. URL https://datatracker.ietf.org/doc/rfc5280/.

[72] The Internet Engineering Task Force. An internet attribute certificate profile for authorization, 2013. URL https://datatracker.ietf.org/doc/rfc5755/?include_text=1.

[73] Telecommunication Standardization Sector of ITU. Security in telecommunications and information technology. Report, International Telecommunication Union, 2004.

[74] Organization for the Advancement of Structured Information Standards. Assertions and protocols for the OASIS security assertion markup language (SAML) v2.0, 2005.

[75] Organization for the Advancement of Structured Information Standards. About us, 2015. URL https://www.oasis-open.org/org.

[76] Walter Isaacson. How Steve Jobs' love of simplicity fueled a design revolution. *Smithsonian Magazine*, 2012. URL http://www.smithsonianmag.com/arts-culture/how-steve-jobs-love-of-simplicity-fueled-a-design-revolution-23868877/?no-ist.

[77] Object Management Group (OMG). Unified modeling language (UML) v2.5, 2015. URL http://www.omg.org/spec/UML/2.5/.

[78] Object Management Group (OMG). Systems modeling language (SysML) v1.4, 2015. URL http://www.omg.org/spec/SysML/.

[79] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014.

[80] International Standards Organization. Industrial automation systems and integration – JT file format specification for 3D visualization, 2012.

[81] Dave Johnson and Steve Speicher. Open Services for Lifecycle Collaboration Core Specification Version 2.0, 2013. URL http://open-services.net/bin/view/Main/OslcCoreSpecification.

[82] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 239: Application protocol: Product life cycle support, 2012.

[83] Timothy D. West and Mark Blackburn. Is digital thread/digital twin affordable? a systemic assessment of the cost of DoD's latest manhattan project. In *Procedia Computer Science*, volume 114, pages 47–56. Elsevier Science Publishers B. V., 2017. doi: 10.1016/j.procs.2017.09.003.

[84] Edward M Kraft. The Air Force Digital Thread/Digital Twin - Life Cycle Integration and Use of Computational and Experimental Knowledge. In *54th AIAA Aerospace Sciences Meeting, 2016, January 4, 2016 - January 8, 2016*. American Institute of Aeronautics and Astronautics Inc, AIAA, San Diego, CA, United states, 2016. doi: doi:10.2514/6.2016-089710.2514/6.2016-0897.

[85] R Wardhani and X Xu. Model-based manufacturing based on STEP AP242. In *12th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pages 1–5, Auckland, New Zealand, 2016. doi: 10.1109/MESA.2016.7587187.

[86] Leonhard Euler. Solutio problematis ad geometriam situs pertinentis. *Commentarii Academiae scientiarum imperialis Petropolitanae*, t.8 (1736-1738):128–140, 1738.

[87] Arthur Cayley. Ii. a memoir on the theory of matrices. *Philosophical Transactions of the Royal Society of London*, 148:17–37, 1858. doi: 10.1098/rstl.1858.0002. URL http://rstl.royalsocietypublishing.org/content/148/17.short.

[88] James Joseph Sylvester. On an application of the new atomic theory to the graphical representation of the invariants and covariants of binary quantics, – with three appendices. *American Journal of Mathematics, Pure and Applied*, 1(1):64–104, 1878.

[89] Frank Harary and John S. Maybee. *Graphs and applications: Proceedings of the first Colorado Symposium on Graph Theory*. John Wiley & Sons Inc, New York, 1985. ISBN 978-0-4718-8772-0.

[90] Robin J. Wilson and John J. Watkins. *Graphs : an introductory approach : a first course in discrete mathematics*. Wiley, New York, 1990. ISBN 978-0-4716-1554-5.

[91] Wai-Kai Chen. *Graph theory and its engineering applications*. World Scientific, River Edge, NJ, 1997. ISBN 978-9-8102-1859-1.

[92] Bolian Liu and Hong-Jian Lai. *Matrices in Combinatorics and Graph Theory*. Springer US, New York, 2000. ISBN 978-1-4419-4834-2.

[93] Jørgen Bang-Jensen and Gregory Gutin. *Digraphs : theory, algorithms, and applications*. Springer, London, 2001. ISBN 978-1-8523-3268-6.

[94] Ravipudi Venkata Rao. *Decision Making in the Manufacturing Environment*. Springer Series in Advanced Manufacturing. Springer London, London, 2007. ISBN 978-1-8462-8818-0. doi: 10.1007/978-1-84628-819-7.

[95] William Thomas Tutte. *Graph theory as I have known it*. Clarendon Press, Oxford, 2012. ISBN 978-0-1996-6055-1.

[96] Narsingh Deo. *Graph theory with applications to engineering & computer science*. Dover Publications, Inc., Mineola, New York, 2016. ISBN 978-0-4868-0793-5.

[97] Jingran Li, Fei Tao, Ying Cheng, and Liangjin Zhao. Big data in product lifecycle management. *International Journal of Advanced Manufacturing Technology*, 81(1-4): 667–684, oct 2015. ISSN 14333015. doi: 10.1007/s00170-015-7151-x. URL http://link.springer.com/10.1007/s00170-015-7151-x.

[98] Oleg Shilovitsky. PLM and data management in 21st century. In *TechSoft3D Tech Talk*, Boston MA, 2013. SlideShare.net. URL https://www.slideshare.net/olegshilovitsky/plm-dm21stcentury.

[99] Thomas D Hedberg Jr, Allison Barnard Feeney, Moneer M Helu, and Jaime A Camelio. Toward a lifecycle information framework and technology in manufacturing. *Journal of Computing and Information Science in Engineering*, 17(2):021010, feb 2017. ISSN 1530-9827. doi: 10.1115/1.4034132.

[100] William Regli, Jarek Rossignac, Vadim Shapiro, and Vijay Srinivasan. The new frontiers in computational modeling of material structures. *Computer-Aided Design*, 77: 73–85, 2016. doi: 10.1016/j.cad.2016.03.002.

[101] Kenneth E. Knight. A descriptive model of the intra-firm innovation process. *The Journal of Business*, 40(4):478–496, 1967. ISSN 00219398. doi: 10.2307/2351630. URL http://www.jstor.org/stable/2351630.

[102] Michael E. Porter. *The competitive advantage of nations*. Free Press, New York, 1990. ISBN 0029253616.

[103] Christopher Freeman and Luc Soete. *The economics of industrial innovation*. MIT Press, Cambridge, Mass., 3rd edition, 1997. ISBN 0262561131 (alk. paper) 0262061953 (alk. paper).

[104] Joseph Tidd and J. R. Bessant. *Managing innovation : integrating technological, market, and organizational change*. Wiley, Hoboken, NJ, 4th edition, 2009. ISBN 9780470998106 (pbk. alk. paper).

[105] Rick M. A. Hollen, Frans A. J. Van Den Bosch, and Henk W. Volberda. The role of management innovation in enabling technological process innovation: An inter-organizational perspective. *European Management Review*, 10(1):35–50, 2013. ISSN 1740-4762. doi: 10.1111/emre.12003. URL http://dx.doi.org/10.1111/emre.12003.

[106] Thomas W. Malnight. Emerging structural patterns within multinational corporations: Toward process-based structures. *The Academy of Management Journal*, 44(6):1187–1210, 2001. ISSN 00014273. doi: 10.2307/3069396. URL http://www.jstor.org/stable/3069396.

[107] Vittorio Chiesa and Federico Frattini. Commercializing technological innovation: Learning from failures in high-tech markets*. *Journal of Product Innovation Management*, 28(4):437–454, 2011. ISSN 1540-5885. doi: 10.1111/j.1540-5885.2011.00818.x.

[108] Melissa A. Schilling. *Strategic management of technological innovation*. McGraw-Hill/Irwin, New York, 2005. ISBN 0072942983 (alk. paper).

[109] Teresa M. Amabile. Creativity and innovation in organizations. *Harvard Business Review*, 1996. URL https://hbr.org/product/creativity-and-innovation-in-organizations/396239-PDF-ENG.

[110] Kurt Lewin, Fritz Heider, and Grace M. Heider. *Principles of topological psychology*. McGraw-Hill publications in psychology. McGraw-Hill book company, inc., New York, London,, 1st edition, 1936.

[111] Martin Hoegl and K. Praveen Parboteeah. Creativity in innovative projects: How teamwork matters. *Journal of Engineering and Technology Management*, 24(1–2): 148–166, 2007. ISSN 0923-4748. doi: 10.1016/j.jengtecman.2007.01.008. URL http://www.sciencedirect.com/science/article/pii/S0923474807000094.

[112] Robert G. Cooper. *Winning at new products : accelerating the process from idea to launch*. Perseus Pub., Cambridge, Mass., 3rd edition, 2001. ISBN 0738204633 (pbk. alk. paper).

[113] Michael L. Tushman and III O Reilly, Charles A. Ambidextrous organizations: Managing evolutionary and revolutionary change. *California Management Review*, 38(4): 8–30, 1996. ISSN 00081256.

[114] Katsuhiko Ogata. *Modern control engineering*. Prentice Hall, Upper Saddle River, NJ, 4th edition, 2002. ISBN 0130609072.

[115] Katsuhiko Ogata. *State space analysis of control systems*. Instrumentation and controls series. Prentice-Hall, Englewood Cliffs, N.J.,, 1967.

[116] Eric Flamholtz. *Effective management control : theory and practice*. Kluwer Academic Publishers, Boston, 1996. ISBN 0792396995 (acid-free paper). URL Publisherdescriptionhttp://www.loc.gov/catdir/enhancements/fy0819/95052283-d.htmlTableofcontentsonlyhttp://www.loc.gov/catdir/enhancements/fy0819/95052283-t.html.

[117] Kim Warren. *Strategic management dynamics*. J. Wiley & Sons, Chichester, West Sussex, England ; Hoboken, NJ, 2008. ISBN 9780470060674.

[118] Francesco Ceresia. A model of goal dynamics in technology-based organizations. *Journal of Engineering and Technology Management*, 28(1–2):49–76, 2011. ISSN 0923-4748. doi: 10.1016/j.jengtecman.2010.12.004.

[119] Peter F. Drucker. *The practice of management*. Harper, New York,, 1st edition, 1954.

[120] Robert Simons. *Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal*. Harvard Business Press, 2013. ISBN 9781422160671.

[121] International Society of Automation. Enterprise-Control System Integration – Part 3: Activity Models of Manufacturing Operations Management, 2013. ANSI/ISA-95.00.03.2013.

[122] Chris Argyris and Donald A. Schön. *Organizational learning*. Addison-Wesley OD series. Addison-Wesley Pub. Co., Reading, Mass., 1978.

[123] Donald G. Reinertsen. *Managing the design factory : a product developer's toolkit*. Free Press, New York, 1997. ISBN 0684839911.

[124] Thomas D Hedberg Jr, Allison Barnard Feeney, and Jaime A Camelio. Towards a diagnostic and prognostic method for knowledge-driven decision making in smart manufacturing technologies. In Azad M Madni, Barry Boehm, Roger G Ghanem, Daniel Erwin, and Marilee J Wheaton, editors, *Disciplinary Convergence in Systems Engineering Research*, chapter 60. Springer International Publishing, 2017. ISBN 978-3-319-62216-3. doi: 10.1007/978-3-319-62217-0.

[125] Robert Simons. *Levers of Control: How Managers Use Innovative Control Systems to Drive Strategic Renewal*. Harvard Business Press, 2013. ISBN 9781422160671.

[126] Joachim Rossberg. Application lifecycle management. In *Pro Visual Studio Team System Application Lifecycle Management*, chapter 2, pages 1–319. Apress, Berkeley, CA, 2009. doi: 10.1007/978-1-4302-1079-5.

[127] John Stark. *Product Lifecycle Management: 21st Century Paradigm for Product Realisation*. Springer-Verlag London Ltd, London, 2 edition, 2011. ISBN 978-0-85729-545-3. doi: 10.1007/978-0-85729-546-0.

[128] Moneer Helu, Alex Joseph, and Thomas Hedberg Jr. A standards-based approach for linking as-planned to as-fabricated product data. *CIRP Annals*, 67(1):487–490, 2018. doi: 10.1016/j.cirp.2018.04.039.

[129] Moneer Helu, Thomas Hedberg Jr, and Allison Barnard Feeney. Reference architecture to integrate heterogeneous manufacturing systems for the digital thread. *CIRP Journal of Manufacturing Science and Technology*, 19:191–195, nov 2017. doi: 10.1016/j.cirpj.2017.04.002.

[130] Tore Dybå and Torgeir Dingsøyr. Empirical studies of agile software development: A systematic review. *Information and Software Technology*, 50(9-10):833–859, aug 2008. doi: 10.1016/j.infsof.2008.01.006.

[131] Çağdaş Üsfekes, Murat Yilmaz, Eray Tuzun, Paul M. Clarke, and Rory V. O'Connor. Examining reward mechanisms for effective usage of application lifecycle management tools. In *Communications in Computer and Information Science*, volume 748, pages 259–268. Springer, 2017. doi: 10.1007/978-3-319-64218-5_21.

[132] Andreas Deuter and Stefano Rizzo. A critical view on PLM/ALM convergence in practice and research. *Procedia Technology*, 26:405–412, 2016. doi: 10.1016/j.protcy.2016.08.052.

[133] Amer Catic and Petter Andersson. Manufacturing experience in a design context enabled by a service oriented plm architecture. In *ASME 2008 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 5, pages 257–265, Brooklyn, New York, 2008. American Society of Mechanical Engineers. doi: 10.1115/DETC2008-49858.

[134] Nenad Ivezic, Boonserm Kulvatunyou, and Vijay Srinivasan. On architecting and composing through-life engineering information services to enable smart manufacturing. *Procedia CIRP*, 22(1):45–52, jan 2014. doi: 10.1016/j.procir.2014.07.004.

[135] Steven J Fenves. *A core product model for representing design information*. National Institute of Standards and Technology, Gaithersburg MD, 2002. doi: 10.6028/NIST.IR.6736.

[136] Steven J Fenves, Sebti Foufou, Conrad Bock, Rachuri Sudarsan, Nicolas Bouillon, and Ram D. Sriram. CPM 2: A revised core product model for representing design information. Technical report, National Institute of Standards and Technology, Gaitherburg MD, 2004.

[137] Tim Berners-Lee. Linked data, 2006. URL http://www.w3.org/DesignIssues/LinkedData.html.

[138] World Wide Web Consortium. Semantic web, 2006. URL https://www.w3.org/standards/semanticweb/.

[139] Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific American*, May, 2001. URL http://www.scientificamerican.com/article/the-semantic-web/.

[140] Jae H. Lee, Steven J. Fenves, Conrad Bock, Hyo-Won Suh, Sudarsan Rachuri, Xenia Fiorentini, and Ram D. Sriram. A semantic product modeling framework and its application to behavior evaluation. *IEEE Transactions on Automation Science and Engineering*, 9(1):110–123, jan 2012. doi: 10.1109/TASE.2011.2165210. URL http://ieeexplore.ieee.org/document/6019021/.

[141] Manas Bajaj, Dirk Zwemer, Rose Yntema, Alex Phung, Amit Kumar, Anshu Dwivedi, and Manoj Waikar. MBSE++ – foundations for extended model-based systems engineering across system lifecycle. In *INCOSE International Symposium*, volume 26, pages 2429–2445, Edinburgh, Scotland, UK, jul 2016. Wiley-Blackwell. doi: 10.1002/j.2334-5837.2016.00304.x.

[142] Manas Bajaj, Jonathan Backhaus, Tim Walden, Manoj Waikar, Dirk Zwemer, Chris Schreiber, Ghassan Issa, and Lockheed Martin. Graph-based digital blueprint for model based engineering of complex systems. In *INCOSE International Symposium*, volume 27, pages 151–169, Adelaide, Australia, jul 2017. Wiley-Blackwell. doi: 10.1002/j.2334-5837.2017.00351.x.

Chapter 4

Conceptual Framework and Technology Innovation[1]

4.1 Introduction

Contextualizing data from the product lifecycle to make design decisions is very difficult. Different data in the product lifecycle is stored in different locations with different people using the data in different ways and in different contexts. The significant difference in data across the lifecycle is the reason why industry anecdotally says, "the lifecycle is starving for information, but drowning in data." A solution is needed to link all the disparate systems of the lifecycle and cultivate information for decision support. Propose is a Lifecycle Information Framework and Technology (LIFT) concept to develop and integrate technology and standards to enable a novel and straightforward product lifecycle management (PLM) implementation that is intelligent, self-learning, and self-aware. The LIFT concept would stretch and/or replace current PLM paradigms with innovation processes and technologies to remove the "silo affect" between organizations. The intent is to create the "Google" for engineering and manufacturing data that supports data curation and information cultivation in an efficient and effective manner. The LIFT concept supports a "data observatory" wherein a user of the PLM system-of-systems would be able to search, discover, and retrieve information from throughout the enterprise when the information is needed. Our viewpoint of a data observatory, synonymous to an astronomical observatory, is a technology that supports the study of engineering and manufacturing phenomena and events through the use of data from the product lifecycle. This paper presents the LIFT concept – a framework for lifecycle information management and the integration of emerging and existing technologies, which together form the basis of a research agenda for a common model to support digital-data curation in manufacturing industries.

This chapter first provides a discussion of the existing technologies and activities that the LIFT concept leverages and integrates in novel ways. Then, the chapter describes the motivation for applying such work to the domain of manufacturing. Then, the proposed LIFT

[1] This chapter was published as an article with the citation: Hedberg Jr, T., Barnard Feeney, A., Helu, M., & Camelio, J. A. (2017). Towards a Lifecycle Information Framework and Technology in Manufacturing. *Journal of Computing and Information Science in Engineering*, 17(2), 021010-021010-021013. doi:10.1115/1.4034132

concept is described. Underlying technologies are further examined. A use case is detailed. Lastly, potential impacts are explored.

4.2 Motivation for Manufacturing

Manufacturing organizations are increasingly using digital engineering artifacts in the product lifecycle. Industry is calling for a digital thread to stitch the phases of the product lifecycle together [1]. PLM theory describes several phases within the product lifecycle. For the purposes of this chapter, the product lifecycle is defined to be the design, analysis, manufacturing, quality assurance, and customer and product support phases. A digital engineering artifact is an object created and managed using primarily software tools (e.g., computer-aided design (CAD), computer-aided manufacturing (CAM)) [2]. model-based definition (MBD) is a type of digital engineering artifact. To manage the data within the software systems, the tools implement proprietary data formats for storing the data. Due to various data format changes aligned historically with new product introductions from the software vendors, industry consortia developed various standard open-data formats (e.g., Jupiter Tesselation (JT) [3], Portable Document Format (PDF) [4] / Product Representation Compact (PRC) [5], Standard for the Exchange of Product Model Data (STEP) [6]). In addition, industry has adopted domain-specific ad-hoc format specifications (e.g., ACIS [7], Stereolithography (STL) [8]) that are published openly by software vendors.

A major challenge for the product lifecycle is the existence of various data format standards, and the existence of little practice standards and no lifecycle information standards. While there are many documented mapping efforts to create interoperability between domain-specific data standards, the majority of the mappings have no way of determining what data and context are required for each phase of the product lifecycle. This leads to information being lost with every data translation – starting with the first translation out of the software tool (e.g., CAD) where the data originated. Moreover, data coming from the authoring CAD system are typically only shape representations [2]. To make sense of the data, the data users in the product lifecycle may also require information about the provenance [9] of the data, feature semantics [10] of the product, and/or activities within an organizational workflow [2].

In addition, the bill of materials (BOM) differs between the various phases of the product lifecycle because downstream functions (e.g., manufacturing, quality, product support) require additional information that engineering does not include in the original BOM. The downstream functions take the BOM coming from engineering and modify it or generate new BOMs to meet functional needs. This has led to creating multiple BOMs with different lifecycle viewpoints. The most commonly found BOMs are the engineering/design BOM (eBOM), the manufacturing BOM (mBOM), and, most recently, the maintenance and support BOM (sBOM). This highlights an interoperability problem similar to the interoperability issues in CAD systems. While there are common elements in each of the BOMs, keeping all the disconnected data flows synchronized is difficult.

The lack of a common-elements model for the product lifecycle is the key driver for the LIFT concept. The Quality Information Framework (QIF) standard [11] started to tip the scales toward common elements. Unfortunately, QIF is not broad enough for all of the information in the entire product lifecycle because the standard was developed to address only the metrology domain-specific issues. Historically, industry using the available commercial PLM solutions [12, 13, 14] suggested managing all of the lifecycle data in a singular (homogeneous) system as a workaround to the lack of a common-elements model.

Estimates [15] point out that data interoperability costs the automotive supply chain over $1 billion[2], but the cost of knowledge transfer and PLM is immeasurable currently. In practice, PLM requires the consideration of many product-data forms beyond the simple inclusion of CAD and BOM information. However, solution providers and industry often conflate singular large-enterprise product-data-management tools and PLM.

The industry's and the commercial PLM vendors' suggestion for a homogeneous system across the enterprise is unrealistic because it is cost prohibitive for small-to-medium enterprise (SME). Further, product-data management (PDM), enterprise resource planning (ERP), manufacturing execution system (MES), and quality management system (QMS) systems were built to solve different problems in the lifecycle. Trying to integrate all of those requirements into a homogenous system would result in a system that is a, "Jack of all trades, master of none [16]." The motivation of this paper is to introduce a concept for a lifecycle information framework that supports a common-elements model of the product lifecycle, which would enable the development and implementation of technology built around the information needs of the product lifecycle while utilizing a heterogeneous system-of-systems.

More recently, the PLM domain has shifted towards "platformization" [17]. The platform concept calls for a foundational infrastructure that represents a PLM operating system. This system would provide a baseline set of capabilities to the end-user. Then, applications, or "Apps," could be plugged into the system to extend the system's capabilities. The platform concept has been successful in the smart-phone domains – Apple iOS(TM) and Google Android(TM) are two examples. Currently, each of the major commercial PLM vendors are developing platform solutions. While, platforms work well in the smart-phone domain, platforms may not be the answer for manufacturing. Much like the single-vendor homogeneous PLM solutions, what platform should industry select? Industry needs a universal plug-and-play solution that would enable native integration of the many systems across the product lifecycle.

[2]Though over 15 years old, the estimate is still relevant. In fact, the authors propose the costs are much more significant than the original estimate due to the increase in the complexity of new products and systems used to design those products.

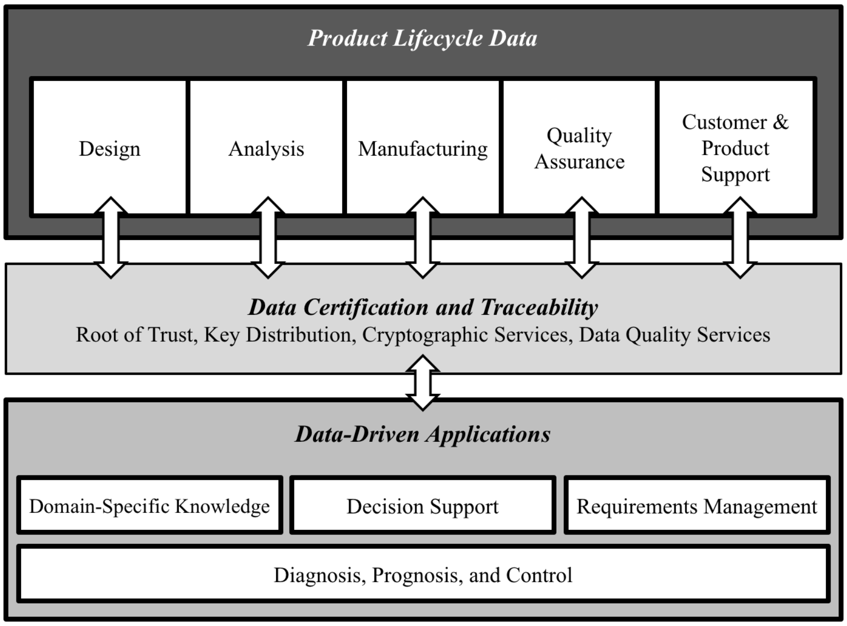


Figure 4.1: Essential supporting structure of the LIFT concept and the relationship of the three layers that make up the framework.

4.3 The Framework

Figure 4.1 presents the structure of the LIFT concept integrated with the product lifecycle. The framework consists of three layers: (1) product-lifecycle data, (2) data certification and traceability, and (3) data-driven applications. This section describes each layer in detail.

Product lifecycle data makes up the first layer of the framework. Recall, the design, analysis, manufacturing, quality assurance, and customer and product support phases define the product lifecycle for the purpose of this paper. The design phase encompasses design processes and functions, which require data be approved by a person with the appropriate authority, certified by providing confirmation the data is what it is declared to be, and then authorization provided for how the data may be used. The analysis phase of the lifecycle analyzes a product using computer-aided engineering (CAE) tools (e.g., simulation-based finite-element analysis (FEA) and computational fluid dynamics (CFD)). The manufacturing phase is where a product is fabricated. The quality assurance phase deals with the inspection and measurement of a product. Lastly, the customer and product support phase manages the end-user support through maintenance services, technical publications, and other support activities. Each of these phases have its own set of data that it keeps for recording-keeping. This data would benefit other phases of the lifecycle too, but linking the data today is difficult and requires significant human capital.

The design development activity is often supported by a design knowledge base. The knowledge base contains meta-data, rules, standards, correlations, and design dictionaries. The knowledge base consists of different types of information transformed into knowledge to support decision making. Those decisions may be manual or automated. Regardless of how the decisions are made, typically the design process must follow some type of knowledge base to be able to design a product.

The product can be a new product introduction, for research and development, or a revision of an existing product. The knowledge base should support the decision-making for all product types based on the knowledge collected over time. However, current industrial knowledge-based methods require manual input and manual interpretation of the knowledge – usually through documents in large policy books or on-line repositories of text-based documents. If a knowledge base has automation, the automation is typically rules-based. In the majority of cases, a human must be able to read documents to extract information from the knowledge base. To make matters worse, there are often multiple domain-specific knowledge bases (e.g., design, manufacturing, quality) throughout the lifecycle and those knowledge bases are not always in agreement with all the lifecycle requirements because they all have different viewpoints.

As evident in the U.S. Department of Defense (DoD) Military Engineering Data Asset Locator System (MEDALS) [18], the amount of human capital required to keep the data repositories up-to-date is significant. The lifecycle cannot afford to deploy the needed amount of human capital to maintain near-dynamic knowledge bases and still deliver products to market. Industry needs a way to discover data relationships and link the data across the lifecycle. This would enable near-real-time dynamic updating of domain-specific knowledge bases in the lifecycle using machine learning and artificial intelligence methods. In the LIFT concept, the only human capital required to enable to automated updating is a data administrator needs to register a domain-specific knowledge base with a "registry" to ensure data is discoverable. Once that registration is complete, the knowledge base would receive near-real-time dynamic updates as users complete their day-to-day activities.

In the middle of the framework is the *Data Certification and Traceability* layer. The data certification and traceability layer supports building trust throughout the product lifecycle. Throughout the whole lifecycle there are different requirements that come in and out of the lifecycle. A lot of those requirements from different phases of the lifecycle often contradict or compete with each other. This raises challenges for the lifecycle to be able to manage all of those requirements, to understand the requirements, and to use the requirements effectively. Misunderstanding and/or not complying with all the requirements leads to distrust of the data.

To enable and ensure trust, the framework needs cryptographic and product-data quality (PDQ) services available through the data certification and traceability layer. The PDQ services would ensure the product data is verified and validated against the multiple sets of requirements that are constraints on the product in the lifecycle.

The PDQ services could interface a *Requirements-Management* application in the *Data-*

Driven Applications layer the framework. The data-driven applications layer would support integrating applications through using plug-and-play methods. Initially, our framework could include applications to support domain-specific knowledge management, decision support, requirements management, diagnosis, prognosis, and control.

A requirements-management application could work closely with a knowledge-management application to ensure all of the lifecycle requirements are captured, understood, and available for reuse. Once the PDQ services complete the verification and validation activities, the cryptographic services embed digital certificates in the product data to create *digital fingerprints* that enable authentication, authorization, and traceability through the lifecycle. The certificates assure the product data is what it says it is (i.e., authentication) and the product data can be used how it is intended to be used (i.e., authorization). Moreover, the certificates support traceability by capturing who did what to whom and when it was done. Overall, the certificates bring seamless authentication, authorization, and traceability to the product data.

Having always up-to-date knowledge bases would support a *Decision Support* application. Working in concert with a knowledge-base application and a requirements-management application, the decision-support application could provide near-real-time feedback to a user as decisions are made. For example, the design knowledge base and requirements manager could build design for manufacturing (DFM) rules based on diagnostic and prognostic data feedback from manufacturing and quality assurance. Those DFM rules would notify the design engineer of tool reach issues based on the data from manufacturing, which would ensure that the design engineer develops a product that manufacturing can produce effectively and efficiently.

Furthermore, with the creation of the QIF standard [11], engineers now have the ability to conduct quality analytics. Quality analytics would allow engineers to look at all the different results, resources, rules, and statistics coming out of a quality organization. The amount of data that can be readily available from the quality-assurance phase of the lifecycle supports the ability to run automated analytics. The analytics can mine data that engineers can turn into design information by applying engineering context to the quality data. Quality analytics is key in generating correlations between the virtual and physical worlds because quality is often the first point in the lifecycle where data is generated from both worlds and compared to each other.

For example, a CAD model could be used during the inspection process to verify the conformance of the physical products. Inspection reports and digital-data sets would be generated that hold valuable data representing the physical product in the cyber-space. This would enable a cyber-physical systems view of the product lifecycle, where the quality data could be used for the purposes of quality assurance, product acceptance, and analyzing what happens throughout the lifecycle. The data-driven application layer of framework would enable the ability to leverage statistical process control, prognosis, and health monitoring methods in novel ways – such as, feed-back and feed-forward between design and supply chain to control the entire product lifecycle.

Customer and product support historically have large amounts of performance data, maintenance records, and customer feedback data that are stored in some location. But a lot of times the data is represented within paper-based records or some disparate database system. This makes getting feedback to the design knowledge base or design engineering role very difficult, if feedback happens at all. The manufacturing functions provide some feedback to engineering, but a lot of the feedback is through ad-hoc discussions. The discussions are started typically because a part cannot be built or manufacturing is having difficulty with a requirement from the design. When formal communication is used between manufacturing and engineering, it is through the use of problem and corrective action reports. But in reality, an e-mail or phone call between the manufacturer and designer are the primary communication methods.

The LIFT concept would enable the input of all the disparate pieces data from the lifecycle-phase silos into a self-learning and self-aware system supported by the data-driven application layer of the framework. Eventually, the system should utilize self-learning algorithms for looking at semantic, syntax, linguistic, and statistic data that comes from all the different data and information repositories of analysis, manufacturing, quality, and customer and product support. The stream of data and information into these self-learning algorithms would enable the system to learn dynamically from streams of data based on product experience. The learning supports near-real-time dynamic updating of design knowledge bases in an effective manner. The effective and traceable information flow, self-learning methods, dynamic updating of the design knowledge base, and real-time decision support form the framework of the LIFT concept.

4.4 The Technology

The backend of the LIFT concept is a derivative of the Handle System – a digital object architecture – developed by the Corporation for National Research Initiatives (CNRI) [19]. The Handle System defines three components: an identifier system, meta-data registries, and digital object repositories [20]. The Handle System architecture enables interoperability for resolution purposes among a varying set of implementations. The Handle System consists of global root servers, local handle servers, clients, and proxy servers – making the system scalable, reliable, and secure for identifier resolution [20]. The purpose of the Handle System is to manage digital objects. A digital object is an abstract entity that has a persistent and unique identifier [20]. The digital object may point a user to a physical, virtual, and/or abstract thing. The most popular implementation of the Handle System is the Digital Object Identifier (DOI®) system [21] supported by the International DOI Foundation [22] and defined in ISO 26324 [23].

The LIFT concept provides a master handling system to act as "traffic cop" of data. The

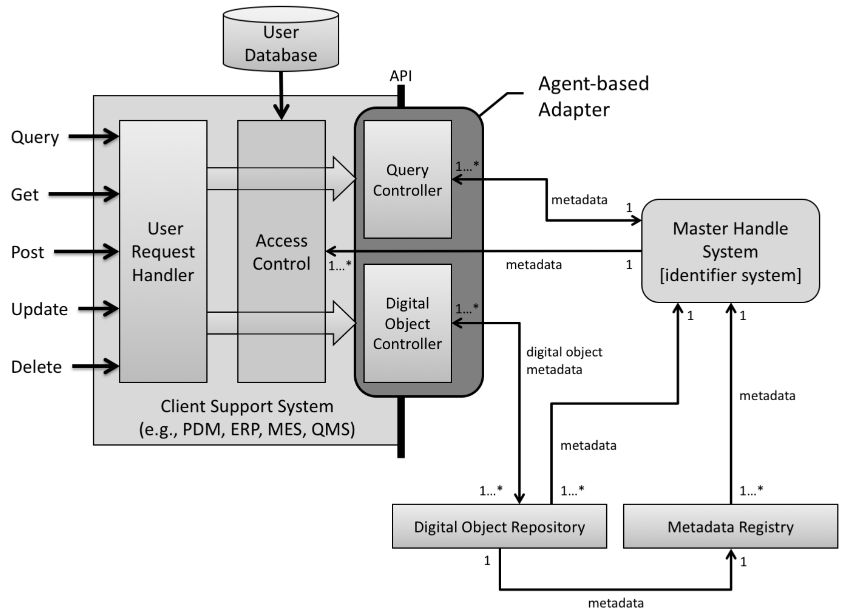


Figure 4.2: Schematic overview of the LIFT concept

LIFT concept builds upon the concept of ISO 26324 [23] to develop a manufacturing-centric product-lifecycle extension to CNRI's Handle System [19]. The potential outcome of this work is the development of a standard for a manufacturing master handling system. The master handling system is the identifier system and resides between all the different databases and/or repositories of data. Figure 4.2 shows the technology architecture of the LIFT concept. The master handling system is an "index-of-indexes" that would understand and be able to inform a user where to go to find particular pieces of information and assist the user with retrieving the appropriate information. The goal of the master handling system is to replace the extensive burden for maintaining links between data that existed in the DoD MEDALS [18] system.

Each piece of information stored in the various repositories across the lifecycle is considered a digital object. Each of the existing databases and/or repositories is considered a digital-object repository. The master handling system does not store or manage the digital objects. The master handling system controls an index of different digital-object repositories and what types of data those digital-object repositories contain. The meta-data registries assist the master handling system with determining the types of data in the digital-objects repositories. In addition, the meta-data registries support access control – an often overlooked kind of meta-data [20] – by controlling who or what can use the digital object and/or how the user can use the digital object. The access control, supported by the data certification and traceability layer of the LIFT framework, enables authentication and authorization capabilities.

Typically, an enterprise utilizes more than one type of data management system that has its own repository and acts as a client support system. These client supports systems are the solutions mentioned earlier at the end of Section 4.2. Engineering and design organizations use the PDM system, manufacturing and supply chain organizations use the MES and ERP systems, and the quality assurance organizations use the QMS systems. Each of those systems already have query, get, post, update, and delete capabilities integrated. The goal of the LIFT concept is to leverage each of the existing client support systems and develop a common element model and integration technologies to support the flow of information between the systems with little or no customization required in the client support systems.

Proposed is an agent-based adapter method for integrating systems through a product lifecycle. Bond and Gasser [24] suggest agents should be used when all of the following are true:

- Data, control, expertise, or resources are inherently distributed.
- The system is naturally regarded as society of autonomous cooperating components.
- The system contains legacy components, which must be made to interact with each other, possibly new software components.

The product lifecycle meets all three elements of Bond's and Gasser's agent-based rationale. Our agent-based adapter method would wrap services to support integrations across the lifecycle. The initial services to target are query control and the digital-object control. Each control service would be a "micro-service" wrapped by the agent-based adapter, which would act as an application programming interface (API) gateway between a client-support system and other product lifecycle systems. Wrapping the control services in an agent-based adapter is an example of a micro-services architecture.

Micro-services architecture requires splitting applications into a set of smaller interconnected services instead of building a singular monolithic applications. A service would implement a set of distinct functionality in our case query or digital-object control. Each functional area of our application could be implemented by its own micro-service. The result is a set of simpler applications and makes it easier to deploy distinct experiences for specific users, devices, or specialized use cases [25].

Using micro-services architecture supports our key goals for the agent-based adapter and integrating various systems across the product lifecycle. Micro-services support scalability and state-less services. Fast plug-and-play would be achievable, which supports low-friction deployment of solutions. The architecture also supports the management of common concerns with infrastructure and framework standards. Deployments only need to worry about business logic. Automated service discovery and routing could be enabled through decoupled services. Flexible integration and registration of systems with the rest of the product lifecycle supports SME with minimal IT resources. Lastly, micro-services would enable a universal Lifecycle Object Identifier (LOI) schema for the handle system by supporting the mapping from vendor-specific and/or proprietary systems to the rest of the product lifecycle.

Most, if not all, enterprise systems have accessible API that allow external systems to interact with the enterprise system using a set of routines, protocols, and/or tools defined by the enterprise systems. Having no standard API for integrating enterprise systems is a major challenge today. The solution providers each develop their own API to be as open or closed as their business models allow. Using the LIFT concept for integrating the enterprise systems, the query controller and digital object controller concept enables pseudo-universal plug and play method for integrating all of the systems in the enterprise.

Additionally, leverage existing service-oriented architecture (SOA) solutions would enable PLM functionality instead of having to develop new SOA capabilities. The existing SOA could be wrapped in the agent-based adapter. The same can be done for existing solution stacks or emerging PLM platforms. Tthose stacks and platforms could be integrated using an coupled agent-based adapter. The agent-based adapter method supports integrating homogeneous systems, platforms, and heterogeneous systems alike – all built with plug-and-play functionality.

The master handling system, through the agent-base adapter, would work in concert with the query controller to enable the building of queries to search for information from the different digital object repositories. The query controller is the link between the user and the master handling system. The query controller would pose a query to the master handling system based on the user's input from the client-support system. The master handling system forms the proper query that is interoperable with all the repositories in the lifecycle. The link between the query controller and master handling system ensures a query across an enterprise is transparent to the user without requiring multiple query inputs from the user. Overall, the queries allow the indexing of all the data that already exists in databases. Then, the indexing supports communicating the data digitally through the lifecycle to the roles and functions that have a need to know based on queries that can be posed by the master handling system.

Once the data or information is discovered in the enterprise, the digital-object controller takes over. The digital-object controller is essentially a data-flow manager. Another goal of the LIFT concept is to not duplicate the data throughout the enterprise. The LIFT concept would work to enable and support data and/or information discovery. The digital-object controller moves data between the source and requesting digital-object repositories as the data is needed. The digital-object controller lets the system clean the data, determine the data quality, and apply the correct context to the data to transform the data into the needed information. This would ensure the data is not just duplicated across the enterprise, but is put to effective and efficient use.

The integration of technology in the LIFT concept forms a semantic web [26] for manufacturing and the entire product lifecycle. The LIFT concept leverages existing ontology research for manufacturing by reusing, expanding, or modifying the ontologies based on the needs for the full product lifecycle. In addition, the extensive research on linked data, described in Section 3.1.4, is reused throughout the master handling system. The query and inference portions of the semantic-web architecture defined by the World Wide Web Consortium (W3C) [26] is a research output of the query controller, digital-object controller, and machine-learning algorithms in the data-driven applications layer that are described in this paper. Lastly, a vertical application of the LIFT concept using semantic web methods is described as a use case example in the next section.

4.5 Use Cases Descriptions

To validate the LIFT concept, the concept was applied to the use cases of the engineering change request (ECR) and dynamic scheduling processes. The ECR and scheduling processes are high-quality use cases because all portions of the product lifecycle have the ability to influence a product change and process improvements. A challenge for industry is to determine when a change is needed in a product or process. Industry struggles with determining when a product or process change is needed because discovering enough information from the product lifecycle is a costly activity.

The LIFT concept supports the information discovery activity with the implementation of the common elements model. Figure 4.3 shows an example of how the common elements model is formed by linking domain-specific element models. The goal of the common element model is to ensure linkage between the minimum amount of information the product lifecycle needs in order to be successful. The LIFT concept does not produce new domain-specific models, but leverages previous research and development to extend existing domain-specific models by linking them together to encapsulate the entire product lifecycle information needs. For example, in Figure 4.3's common elements model, information may flow between the shaded node in the design elements model to the shaded nodes in the manufacturing and quality elements models. The links between the domain-specific models enables data with context to flow between product lifecycle phases and roles by allowing access to various cross-domain nodes required by each phase and role.

Figure 4.4 shows a hierarchal model of the data flow in ECR and dynamic scheduling processes. At the top of the hierarchal model are the lifecycle processes (e.g., design, manufacturing, quality assurance). As you move down the model, there is an abstraction of data between the process layer and the product data, common elements model layer, and the decision layer. The design, manufacturing, and quality assurance activities are part of a process layer. In the product data layer exists the virtual and physical data related to the product. The product definitions (e.g., three-dimensional (3D) models), process monitoring data (e.g., MTConnect [27] output), and quality measurement data (e.g., QIF results files) are examples of the type of data in the product data layer. The next layer is the common elements model layer, which is where data combined with context to generate actionable information. That last layer takes into account knowledge built upon the linked data and information flows through the product lifecycle and supports the actual decisions for ECR and dynamic scheduling processes.

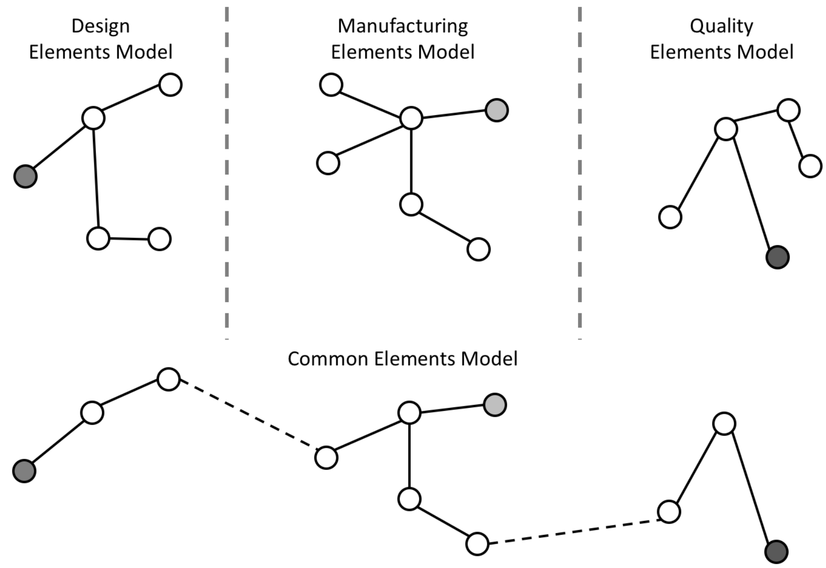


Figure 4.3: Example Common Elements Model formed by linking domain-specific information models

Thus, the data flows out of the process layer into the common elements model by applying the appropriate domain-specific context to the data from each activity. This is the transformation of data into information. That information can be analyzed upon which to build knowledge. The gained knowledge would support decisions that need to be made in the product lifecycle.

The LIFT concept is implemented in all layers of the ECR and scheduling processes shown in Figure 4.4. The common elements model would act as a universal information communicator in the framework portion of the LIFT concept. In the process level, model-based enterprise (MBE) recommended practices are implemented to ensure an effective level of product-data quality through the activities. However, the majority of the LIFT concept is implemented in product data and decision layers of the use cases. The data-driven application layer of the LIFT framework is the link between all the process activities and their associated product data. The applications would leverage various product definitions (e.g., CAD models), MT-Connect [27] output, and QIF [11] to implement linked data methods. The domain-specific knowledge base update engines, decision support engines, and requirements management engines from the LIFT framework are enabled in the decision layer of the ECR use case.

Figure 4.5 describes an example process for automating the ECR process. The data-driven application layer of the LIFT framework interfaces with the knowledge-base engines, decision-support engines, and the requirements-management engines to generate and/or discover information that supports automating the ECR process. The first step in the automated

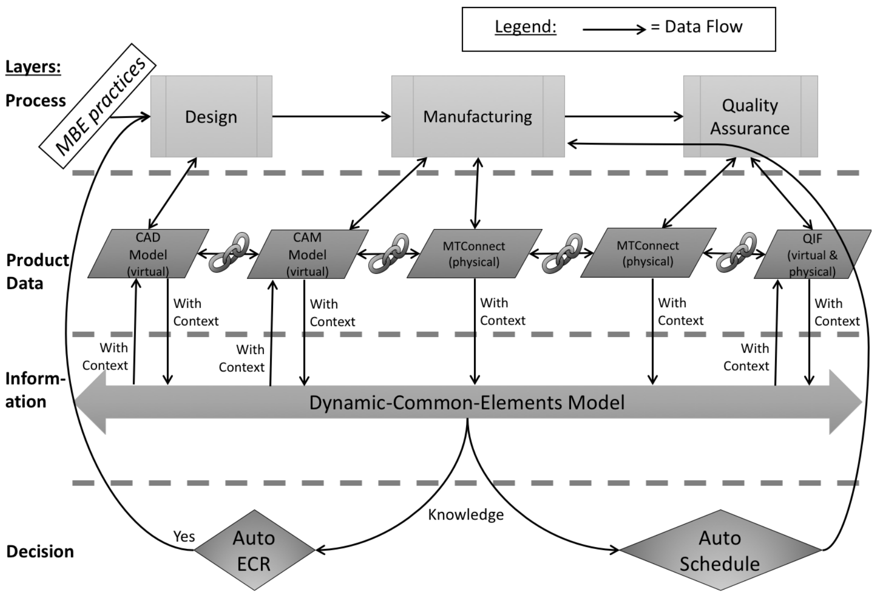


Figure 4.4: Example data flow for engineering change requests to design and automated / dynamic scheduling to manufacturing. Each dotted line, represents a different abstraction layer. Here, is an illustration of how data exchange manifests itself. First, as data is collected out of the process layer, context is added to the data layer to generate the information layer. Then, knowledge is built upon the information layer to support decision making. In this case, deciding when an engineering-change request is required or to determine dynamic scheduling within manufacturing.

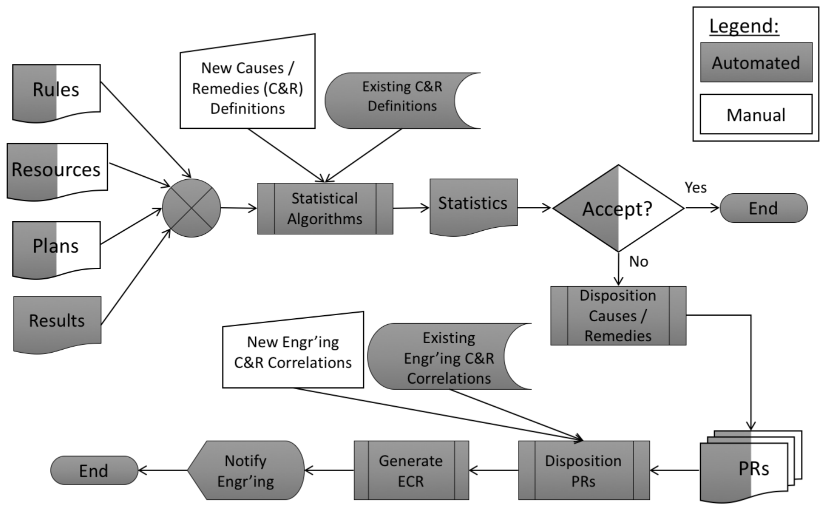


Figure 4.5: Example process for Automated Engineering Change Requests to show how data is first aggregated together, statistical anomalies detected, engineering-design problems discovered, and engineering changes requests generated.

ECR process is to aggregate all the virtual-world and physical-world data.

In Figure 4.5, the lifecycle (e.g., design, manufacturing, and quality) rules, resources, plans, and results are aggregated and links are established. In this use case, only statistical algorithms are considered to simplify the research task, but semantic, syntactic, linguistic, as well as statistics methods must be utilized to automate the ECR decision-making process fully. Aggregating and linking all the lifecycle data supports the identification of product issues. The cause of the issues must be understood and a remedy identified.

Part 8 (Statistics) of QIF [28], describes a schema for defining cause and remedy (C&R) databases. Using the schema could enable supervised machine-learning to automatically identify causes of product issues and apply a recommended remedy. When a product issue is identified and the system cannot apply a cause and remedy definition, a human would be notified to teach the machine-learning system the cause and recommend remedy for the identified product issue. The automated C&R process could alleviate the human-resources burden on the supply chain and product cycle-time by only bringing the human in the loop when the machine-learning system needs supervision. Once a C&R is assigned to a product issue, the decision-support engines could make a determination on the acceptance of the product.

The requirements-management engines would support decision-support engines in making the acceptance decision. If a product is not accepted, the reason for non-acceptance needs to be understood. Therefore, the automated ECR system would need to analyze the C&R assignment and other product lifecycle information to develop a potential problem report.

If enough problem reports are generated for a particular problem, the problem needs to be dispositioned and correlated to a phase of the product lifecycle (e.g., design, manufacturing, quality assurance). In the use case described in this paper, our work is demonstrating statistical process control on the entire product lifecycle by correlating the problem back to the appropriate lifecycle phase. The information used for the development of the problem reports is captured and fed back to the appropriate knowledge bases to ensure that real-time data is accurate and available to decision makers. Just like the C&R definitions databases, C&R correlation databases are created to assist in the supervised machine-learning. The study is testing the use of design engineering C&R correlations. When a problem is correlated back to a design engineering issue, the problem would be dispositioned and an ECR generated. The final step in the process would be to notify engineering of the ECR.

The National Institute of Standards and Technology (NIST) operates the Smart Manufacturing Systems Test Bed (SMS Test Bed) [29] to support validating use cases and concepts. The SMS Test Bed is comprised of two components: (1) a computer-aided technologies (CAx) laboratory and (2) a manufacturing laboratory. Both components focus primarily on machining processes, but the goal of the test bed is to generate a systems-level perspective applicable beyond specific manufacturing process domains. The CAx laboratory utilizes various design, manufacturing, inspection, data management, and verification and validation tools. The manufacturing laboratory various milling and turning machine tools, inspection equipment, production management systems, and data acquisition tools.

Each layer of the LIFT concept was validated. Continue testing of the LIFT concept and use cases would further validate the full integration of the LIFT concept. The next step is to run single-blind experiments. The idea is to develop a product definition with a known design defect that causes manufacturing and quality deficiencies. The system designers, and therefore the system algorithms, would not have knowledge of the design defect. The product definition would be generated in CAD systems and translated to STEP Application Protocol (AP) 242 [6] for exchange with the manufacturing and quality functions. This includes collecting manufacturing and quality data using the MTConnect [27] and QIF [11] standards while producing the product with the known design defect in the manufacturing laboratory of the SMS Test Bed. Then, aggregate the data using a process like the one outlined in Figure 4.5. If the system is able to automatically detect the defect and deficiencies and correlate them back to engineering, then evidence shows it is feasible to automate engineering change request generation and supported our use case.

4.6 Potential Impacts

The Office of Science and Technology Policy (OSTP) directed Federal agencies with research and development expenditures to develop a plan to support public access to data and results of research funded by the Federal Government [30]. The purpose of the directive was to help Federal agencies comply with the America COMPETES Reauthorization Act of 2010 [31]. In addition, various science and technology communities [32, 33, 34, 35] provide guidance to managing and publishing domain-specific data.

However, the engineering and manufacturing domain may not share the same view of data sharing as the physical sciences domain. Due to intellectual property (IP) and data-security concerns, most manufacturing data remains behind lock and key – much less discoverable. The creation and rollout of the National Network of Manufacturing Innovation (NNMI) would help change the current thinking for the management of manufacturing data. The NNMI aims to support the development of a United States infrastructure for manufacturing research [36] and may be subject to the OSTP mandate.

For the NNMI to be successful, a data management plan that is acceptable to all parties and consistent across the network must be available. The LIFT concept provides a common data management plan that supports data discovery activities and manages the authentication and authorization control around the data. The reference infrastructure from the LIFT concept and demonstrated in the NIST SMS Test Bed [29] provides deployable data management guidance to the various institutes that make up the NNMI. By deploying a common infrastructure across the NNMI, the manufacturing data generated in the network may be discoverable across the network. Making the data in the NNMI discoverable and/or available via a common infrastructure benefits the public, industry, academia by enabling researchers to understand and exploit existing manufacturing data for innovation and discovery – in much the same way witnessed in the forecasting industry with open weather data or the biotechnology sector with public access to genome sequences [30]. The exploitation of manufacturing data would help accelerate industry into the the third industrial revolution [37] where data becomes the commodity through the digitization of manufacturing.

The LIFT concept also provides a reference architecture for a manufacturing application of semantic web technologies – further driving the commoditization of data. This enables industry to build a grid of product data across the product lifecycle to support real-time decision support and knowledge building. From organizational learning theory, people have mental maps with regard to how to act in situations [38]. The mental maps people form are driven by people's experiences in previous situations and involves the way people plan, implement, and review their actions [38]. The maps guide people's actions rather than the theories the people espouse explicitly causing a split between theory and actions [39]. In control theory, the mental map is a single-loop control system.

Recall, Figure 3.7 shows a single-loop control system and a double-loop control system. Single-loop control is the repeated attempt at same problem with no variation of method and without ever questioning the goal [40]. Whereas, double-loop control is after attempting to achieve a goal on different occasions, the goal is modified in light of experience or – in the worst case – the goal is rejected all together [40]. Argyris [39] argues double-loop control can be applied to organizations to break the pattern of, colloquially, “this is the way we’ve always done it.” Argyris [39] calls that “double-loop learning.” The LIFT concept has the potential to provide the information needed to reduce the double-loop learning theory to practice and would provide effectively a complete perspective of the product-lifecycle impact of a decision. The LIFT concept would put information into the hands of the roles and functions that have a need to know for decision support and knowledge-management purposes. The LIFT concept would support the cost and quality influencers in the product lifecycle by educating them through double-loop learning on how to make designs better and more producible, which may lead to removing the disconnect between theory and action.

A more far-reaching impact of the LIFT concept is the ability to deploy prognosis, health monitoring, and control methods to all aspects of the product lifecycle (e.g., the product, operations, activities) and not just to the product as prognosis and health monitoring (PHM) is applied historically by industry. The use case described in Section 4.5 only scratches the surface of capabilities enabled by the LIFT concept. The Smart Manufacturing Operations Planning and Control (SMOPAC) program [41] at the NIST is investigating how the LIFT concept applies to all aspects of the product lifecycle from the perspective of the manufacturing domain. The SMOPAC program is investigating how prognostic feed-back and feed-forward information can be applied to the design operations, machine tool, manufacturing operations, quality-inspection tool, quality-assurance operations, and product-support operations activities in the product lifecycle by ingesting manufacturing data and information using the LIFT concept. The goal is to accelerate the benefits and impacts for the public, industry, and academia by maturing the LIFT concept through reference infrastructure to demonstrate linking of all the data in the product lifecycle similar to the method describing the links between the process and product data layers in Figure 4.4 from Section 4.5.

4.7 Conclusion

The framework and technology together makes up the LIFT concept. This paper provided an overview of the emerging information framework and needed technology infrastructure to support semantic product lifecycle management, showed how the framework and technology relates to a real engineering problem and process, and suggested additional research topics to mature the concept. This paper discussed previous and related research to describe how the LIFT concept is aiming not to reinvent the wheel, but leverage and expand the work from other domains to springboard success within the manufacturing domain. This paper used the engineering change request process to illustrate the potential applications of the dynamic knowledge base, decision-support engine, and requirements-management engine components of the framework. The discussion throughout the paper highlighted areas of research that need expansion and questions that need to be solved before the concept can realize its full impact.

A few research questions remain. These questions pertain to calculating uncertainty and variation in the product lifecycle, semantics and domain-specific knowledge, and collaboration and interoperability. Uncertainty quantification is understood at the domain-specific levels, but how do those uncertainties aggregate up into a total uncertainty understanding for the entire product lifecycle? Also, what are the modes of variation in the product lifecycle and where are those models most likely to occur? To support applying control theory to the complete product lifecycle, one must be able to understand how uncertainty and variation across the lifecycle relates. One must know how to predict the uncertainties of lifecycle phases, how those aggregate together, and how to identify where variation could be introduced. From there one must determine how to minimize the uncertainties and variation through decision-support systems and selection mechanisms.

In addition, the LIFT concept can be successful only if semantic links are generated between data, which would support computer-processable knowledge bases. The problem with defining semantics is the definer often requires domain-specific knowledge of the things that the semantics are defining. Therefore, generating links and relationships between domain-specific data and information may be domain-specific itself. The question remains, how would semantic links between data be generated using automated methods?

Lastly, intra-domain (e.g., design to design) and inter-domain (e.g., design to manufacturing) interoperability must be increased. Studies [1, 42, 43] show communicating product data using model-based methods is feasible, but barriers in standards and interoperability remain. For MBE and our concept to be successful, closing the interoperability gaps is paramount.

In closing, the research currently suggests the LIFT concept could manage all the information within the lifecycle by supporting linkages between currently disconnected information. The LIFT concept could also support organizational learning and support the removal of inter-organizational socio-technical barriers. While additional research is required, the outcome of the LIFT concept is a novel solution to the problem that the lifecycle is drowning in data, but starving for information.

Chapter Bibliography

[1] Thomas D. Hedberg Jr, Joshua Lubell, Lyle Fischer, Larry Maggiano, and Allison Barnard Feeney. Testing the digital thread in support of model-based manufacturing and inspection. *Journal of Computing and Information Science in Engineering*, 2016. ISSN 1530-9827. doi: 10.1115/1.4032697. URL `http://dx.doi.org/10.1115/1.4032697`.

[2] William C. Regli, Joseph B. Kopena, and Michael Grauer. On the long-term retention of geometry-centric digital engineering artifacts. *Computer-Aided Design*, 43(7):820–

837, 2011. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/j.cad.2010.11.012. URL http://www.sciencedirect.com/science/article/pii/S0010448510002228.

[3] International Standards Organization. Industrial automation systems and integration – JT file format specification for 3D visualization, 2012.

[4] International Standards Organization. Document management – portable document format – part 1: PDF 1.7, 2008.

[5] International Standards Organization. Document management – 3D use of product representation compact (PRC) format – part 1: PRC 10001, 2014.

[6] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014.

[7] Spatial Corporation. ACIS R25, July 2014. Proprietary Specification.

[8] Marshall Burns. *The StL Format*, pages xxv, 369 p. PTR Prentice Hall, Englewood Cliffs, N.J., 1993. ISBN 0131194623.

[9] Sudha Ram and Jun Liu. *Understanding the Semantics of Data Provenance to Support Active Conceptual Modeling*, volume 4512 of *Lecture Notes in Computer Science*, book section 3, pages 17–29. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-77502-7. doi: 10.1007/978-3-540-77503-4_3. URL http://dx.doi.org/10.1007/978-3-540-77503-4_3.

[10] R. Bidarra and W. F. Bronsvoort. Semantic feature modeling. *Computer-Aided Design*, 32(3):201–225, 2000. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/S0010-4485(99)00090-1. URL http://www.sciencedirect.com/science/article/pii/S0010448599000901.

[11] Dimensional Metrology Standards Consortium. Part 1: Overview and fundamental principles in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014. URL http://www.qifstandards.org.

[12] Dassault Systemes. Achieving better, faster, smarter innovation. Waltham MA, 2012. URL http://www.3ds.com/fileadmin/PRODUCTS-SERVICES/ENOVIA/Resources-center/Achieving-Better-Smarter-Innovation-WP.pdf.

[13] PTC Inc. Windchill (R): Managing the complete product lifecycle – from concept to service. Needham MA, 2014. URL http://www.ptc.com/product/windchill/new.

[14] Siemens PLM. Teamcenter: Simplifying plm. Plano TX, 2014. URL http://www.plm.automation.siemens.com/en_us/products/teamcenter/#lightview%26uri=tcm:1023-79817%26title=Teamcenter-Brochure-4680%26docType=.pdf.

[15] Gregory Tassey, Smita B Brunnermeier, and Sheila A Martin. Interoperability cost analysis of the us automotive supply chain. Report 7007-03, Research Triangle Institute, 1999.

[16] Gregory Y. Titelman. *Random House Dictionary of Popular Proverbs and Sayings*. Random House, New York, 1996.

[17] Peter A. Bilello. The next step in PLM's evolution: Platformization. *Product Design and Develoment*, 2014-12-16 2014. URL http://www.pddnet.com/article/2014/12/next-step-plms-evolution-platformization.

[18] Logistics Information Service. Military engineering data asset locator system (MEDALS). Report, Defense Logistics Agency, 2014. URL https://www.dlis.dla.mil/medals/pdfs/MEDALSHomepageOverview.pdf.

[19] Robert Kahn and Robert Wilensky. A framework for distributed digital object services. *International Journal on Digital Libraries*, 6(2):115–123, 2006. ISSN 1432-5012. doi: 10.1007/s00799-005-0128-x.

[20] Sean Reilly and Robert Tupelo-Schneck. Digital object repository server: A component of the digital object architecture. *D-Lib Magazine*, 16(1/2), January / February 2010. doi: doi:10.1045/january2010-reilly.

[21] International DOI Foundation. DOI Handbook. Report, International DOI Foundation, 2012. URL http://www.doi.org/hb.html.

[22] International DOI Foundation. International DOI Foundation Members, June 2014. URL http://www.doi.org/idf-member-list.html.

[23] International Standards Organization. Information and documentation – digital object identifier system, 2012.

[24] Alan H. Bond and Leslie George Gasser. *Readings in distributed artificial intelligence*. M. Kaufmann, San Mateo, Calif., 1988. ISBN 093461363X.

[25] Chris Richardson. Introduction to microservices, 2015-05-19 2015. URL https://www.nginx.com/blog/introduction-to-microservices/.

[26] World Wide Web Consortium. Semantic web, 2006. URL https://www.w3.org/standards/semanticweb/.

[27] MTConnect Institute. Mtconnect standard, 2014. URL http://www.mtconnect.org/media/39542/mtc_part_1_overview_v1.3.pdf.

[28] Dimensional Metrology Standards Consortium. Part 8: QIF statistics information model and XML schema files in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014.

[29] Moneer Helu and Thomas Hedberg Jr. Enabling smart manufacturing research and development using a product lifecycle test bed. *Procedia Manufacturing*, 1:86–97, 2015. ISSN 2351-9789. doi: http://dx.doi.org/10.1016/j.promfg.2015.09.066.

[30] John P. Holdren. MEMO for increasing access to the results of federally funded scientific research. Report, Executive Office of the President, Office of Science and Technology Policy, 2013. URL http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf.

[31] US Law. America COMPETES Reauthorization Act of 2010, 2011. URL http://www.gpo.gov/fdsys/pkg/PLAW-111publ358/pdf/PLAW-111publ358.pdf.

[32] Consortium for Ocean Leadership. Final network design. Report, Ocean Observatories Initiative, 2010. URL oceanleadership.org/wp-acontent/uploads/2009/04/1101-00000_FND_OOI_2010-04-22_ver_2-06_public1.pdf.

[33] Open Data Institute. Implementation plan 2102 and beyond. Report, Open Data Institute, 2012. URL http://theodi.org/about-us.

[34] National Aeronautics and Space Administration. Physical science information: History and purpose, 2014. URL http://psi.nasa.gov/History.aspx.

[35] EarthCube. Earthcube strategic science plan. Report, National Science Foundation, 2015. URL http://earthcube.org/document/earthcube-strategic-science-plan.

[36] U.S. Department of Commerce. Three takeaways from national network for manufacturing innovation (NNMI) day, 2014. URL http://www.commerce.gov/blog/2014/09/22/three-takeaways-national-network-manufacturing-innovation-nnmi-day.

[37] Economist. Manufacturing: The third industrial revolution. *The Economist*, page 14, 2012. URL http://www.economist.com/node/21553017.

[38] Chris Argyris and Donald A. Schön. *Organizational learning*. Addison-Wesley OD series. Addison-Wesley Pub. Co., Reading, Mass., 1978.

[39] Chris Argyris. *Teaching smart people how to learn*. Harvard business review classics series. Harvard Business Press, Boston, Mass., 2008. ISBN 9781422126004.

[40] Donald G. Reinertsen. *Managing the design factory : a product developer's toolkit*. Free Press, New York, 1997. ISBN 0684839911.

[41] Allison Barnard Feeney and Brian Weiss. Smart manufacturing operations planning and control program, 2014. URL http://www.nist.gov/el/msid/syseng/smopc.cfm.

[42] Kevin Ficher, Phil Rosche, and Asa Trainer. Investigating the impact of standards-based interoperability for design to manufacturing and quality in the supply chain. Report, National Institute of Standard and Technology, 2015. URL http://www.nist.gov/manuscript-publication-search.cfm?pub_id=920033. NIST GCR 15-1009.

[43] Asa Trainer, Thomas Hedberg Jr, Allison Barnard Feeney, Kevin Fischer, and Phil Rosche. Gaps analysis of integrating product design, manufacturing, and quality data in the supply chain using model-based definition, 2016. 2016 ASME Manufacturing Science and Engineering Conference.

Chapter 5

Linking Product Lifecycle Data Dynamically[1]

5.1 Introduction

Between 1998 and 2015, U.S. manufacturing productivity grew three times faster than the service economy [1]. While manufacturing exhibited growth and success, significant opportunity remains. For design through production portion of the product lifecycle, one study found that simply transitioning from paper-based processes to (digital) model-based processes would achieve an approximate 75 percent reduction in cycle-time [2]. Further, enhanced sensing and monitoring, seamless transmission of digital information, and advances in analyzing data and trends would save manufacturers $30 Billion annually [3].

But industry is also approaching the fundamental limits of what its people, tools, and processes can manage. The challenges with managing manufacturing-related data are well understood [4, 5, 6, 7]. Further, data, system, and viewpoint interoperability is an increasing challenge for industry [7, 8, 9, 10, 11]. Industry needs connected systems and linked-data federated across enterprises. Point-to-Point interoperability (e.g., file-based data translation) is no longer enough. Industry must stop thinking about data interoperability through mapping exercises and instead focus on domain and interface interoperability.

Domain interoperability (e.g., design to manufacturing, design to quality) requires a normalized method for accessing and contextualizing data at different points of the product lifecycle. Often the focus of interoperability has been confined to the formats in which the data is stored and not the semantics. Focusing on the information for the "thing" being represented in the data would help industry keep more focus on solving problems for the thing than focusing on communication and data exchange. Further, actors in industry must also consider the interfaces, outputs, and inputs, on the boundaries of their domains. Standard interfaces between domains must be developed and understood to support efficient flow of required information through the product lifecycle. This effective communication of information brings with it an almost $8 Billion return-on-investment annual opportunity [3].

[1] This chapter was submitted as an article with the citation: Hedberg Jr, T., Bajaj, M., & Camelio, J. A. (In Review). Using Graphs to Link Data Across the Product Lifecycle for Enabling Smart Manufacturing Digital Threads. *Journal of Computing and Information Science in Engineering*

Context varies based on the phase of the lifecycle (e.g., design, manufacturing, quality). Each phase of the product lifecycle has different viewpoints and concerns, which lead to different levels of abstraction in modeling and simulation [12, 13]. In addition, context varies based on the level of interaction with data (e.g., systems, operations, enterprises) [11]. The various viewpoints lead to information models and systems being developed for a specific purpose, which results in different information models across the product lifecycle to look at the same data in different ways. Thus, geometry and manufacturing specification is not enough to define products – behavioral and contextual definitions are required too. Furthermore, all three aspects must be generated, documented, and communicated using an agile and dynamic method.

Our standards-based, linked-data approach provides seamless traceability across the product lifecycle, enables high-quality manufacturing, and supports enterprise knowledge reuse. Traceability is import to industry because one must know the provenance of data and/or parts to ensure that those things are trustworthy. Seamless traceability must be supported between the systems, designs, manufacturing operations, and maintenance of products. High-quality manufacturing remains a goal of industry because industry wants to make parts faster, cheaper, and better.

Last, enterprise knowledge reuse supports industry's need in retaining and generating knowledge regardless of what human resources are available. People come and go in organizations, but the knowledge must remain. The goals can be achieved by connecting data across the enterprise to spin a digital thread. Different contextual models can be generated as information moves across disciplines. Also, tracking changes as well as comparing, synchronizing, and repairing connections are topics of interest related to linking data across enterprises. Achieving the goal of a standards-based linked-data approach for distributed, smart manufacturing is the major contribution of the work presented here.

This chapter, in Section 5.2, describes a methodology for linking and tracing data throughout the product lifecycle. Section 5.2.1 specifically addresses the information requirements and architecture proposed for making connections across enterprises to form smart manufacturing digital threads. Further, Section 5.2.2 proposes a method for ensuring persistent global identification. Section 5.3 presents a case study to demonstrate the method applied to an information round-trip between design, manufacturing, and quality. Before concluding, Section 5.4 will discuss generating connections dynamically, forming frequently asked queries enabled by the method, contextualizing graph-based viewpoints, and knowledge generation.

5.2 Information Model and Architecture

A federated digital thread includes artifacts originating from different discipline, tool, and repository ecosystems, such as product lifecycle management (PLM) or application lifecycle management (ALM) systems, in the design-manufacturing-supply-chain network. Although

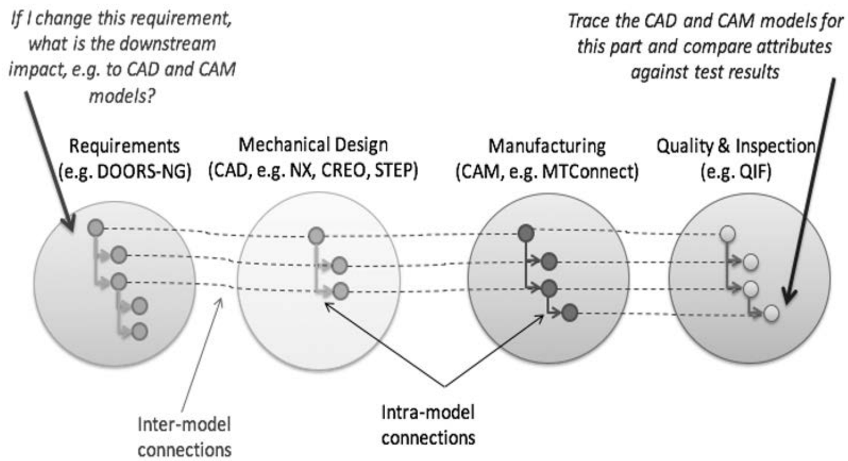


Figure 5.1: A representative example of a digital thread for manufacturing [from [14, 15, 16]].

each repository provides tools to manage the artifacts originating in that repository (e.g., versioning, configuration management, verification and validation), it is the curation of artifacts originating from different repositories that poses a challenge.

Consider a representative subset of a digital thread (See Figure 5.1), which connects artifacts originating in four different repositories: (1) product requirements originating in a requirements management system, (2) mechanical design models originating from a product-data management (PDM) system, (3) MTConnect and other computer-aided manufacturing (CAM) models originating from manufacturing process-planning tools, and (4) quality inspection reports in Quality Information Framework (QIF) originating from a quality management system (QMS). Even though the artifacts originating in a given repository may be seamlessly linked to artifacts originating in the same repository (intra-model connections), it is the inter-model connections between the artifacts across the repositories that enable a federated digital thread. Both intra- and inter-model connections are necessary to traverse and query a graph-based digital thread. Two sample queries shown in Figure 5.1 are:

- If a product requirement changes, can one assess the impact of the change downstream to mechanical/electrical design and manufacturing process plans? The impact may be measured in terms of time, resources, and cost to affect the change.

- If a part fails during operation, can one trace upstream to the mechanical/electrical design and product requirements?

Using the idea of a federated digital thread, artifacts can be connected across entire enterprises. Making connections across enterprises is about abstracting up to a higher level to solve problems and make decisions. The connections must be made using a technology-agnostic approach. Technologies change over time, but the information needs do not. Therefore, the method for making connections must use non-changeable attributes of the artifacts being linked. Some examples of these attributes are location, ownership, or any other attribute that has the possibility of changing without changing the identity of the referent. The goal is to ensure persistent connections regardless of how artifact attributes may change.

Kahn and Wilensky [17] purposed a framework for distributed digital object services. An original motivation for the framework was the need to identify and retrieve information over long periods of time (e.g., tens of years, hundreds of years). Therefore, persistence was a critical design requirement. While Kahn's and Wilensky's framework originally addresses digital objects, the manufacturing sector requires an approach that can manage artifacts that are digital (i.e., cyber) or physical. Connecting only digital objects is not enough for industry because it must also include the connections to the physical world during decision making and problem solving (e.g., traceability analysis, accident investigation). Therefore, an extension to Kahn's and Wilensky's work is proposed. Section 5.2.1 provides our extended digital object architecture that encompasses the Kahn and Wilensky framework and Section 5.2.2 addresses persistent global identification in the context of manufacturing-specific intellectual property (IP) and data-rights issues.

5.2.1 Lifecycle Handler System

Starting the with technology schematic (See Figure 4.2) presented in Section 4.4, the architecture developed for making connections across manufacturing enterprises is shown in Figure 5.2. The architecture forms what the Lifecycle Handler System (LHS). The LHS includes the global handle registry and local handle services from Kahn and Wilensky [17], but also adds client support systems, local graph databases, and agent-based adapters. The LHS enables the ability to expose the digital thread as a set of services so that higher-level analysis and verification applications can be built and deployed for teams across the product lifecycle (e.g., design, manufacturing, and operation).

The LHS system leverages the Handle System [19] to connect to the global handle registry and deploy local handle services. The Handle System was selected as a starting point for our work here because it was developed around the question of how best to connect, track, and access information stored at locations not always known [17]. Further, the underlying architecture of the Handle System accounted for IP issues as a critical component of the undertaking. Last, backed by the ISO 26324 standard [20], the Distributed Object Identifier (DOI) system[2] is based on the Handle System and the International DOI Foundation has expanded the scope of a handle to be a digital identifier of an object, which they define as "thing: physical, digital, or abstract," [21]. The DOI system serves primarily the media and publication sectors and has approximately 175 million DOI names assigned to date with over 5 billion DOI resolutions per year.

In the LHS, handles are generated and managed in accordance with RFC 3650 [19], RFC

[2] http://www.doi.org

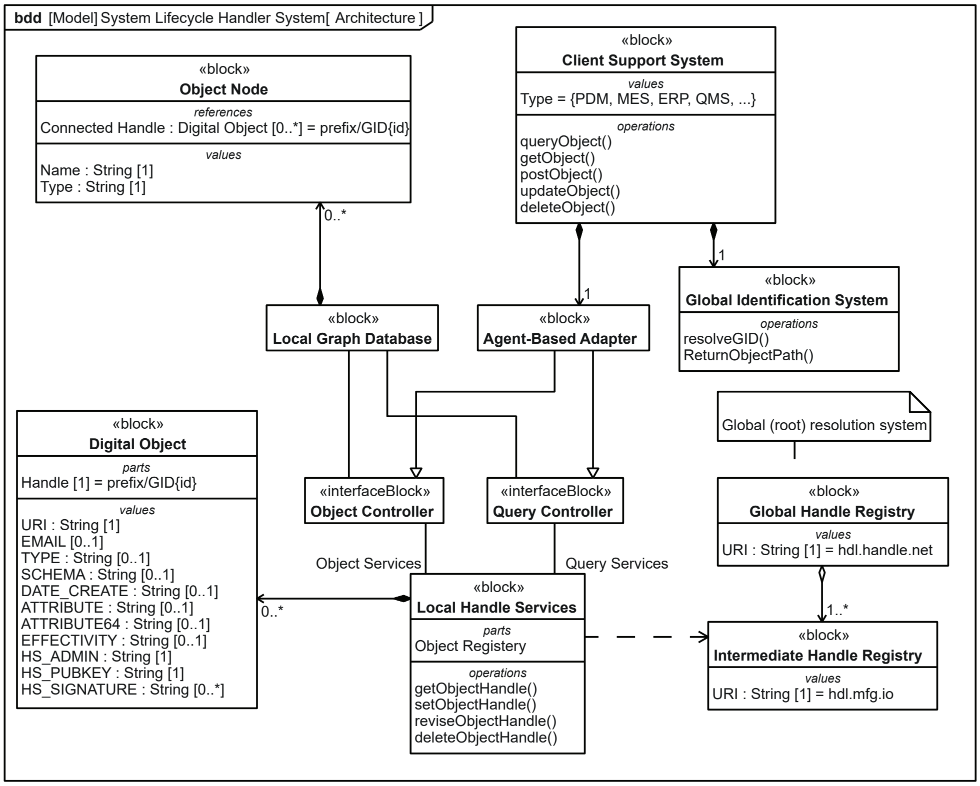


Figure 5.2: An architecture for making connections across enterprises based on the Lifecycle Information Framework and Technology (LIFT) concept [18].

3651 [22], and RFC 3652 [23]. A handle is composed of a naming authority and local name. The handle syntax is shown in Listing 5.1. Figure 5.3 provides an overview of the process defined by RFC 3650 for resolving handles from the global handle registry to the local handle services. The client queries the global handle registry to determine which local handle service manages the handle's prefix. Then, the client queries that local handle services to retrieve the information about the handle. Finally, the client processes the returned information in accordance with the requested action.

An agent-based adapter composed of micro-services for query and object control is attached to client support systems in the LHS. The adapter tracks activity within the client support systems and captures links between artifacts. The adapter stores the handles of artifacts as nodes in a local graph database. The handles of each linked connection are also captured for the nodes in the database. For example, if a CAM model is generated using a portion of a computer-aided design (CAD) model, a node is generated for both models, the handles of each model is captured, and a directed or undirected edge is generated between the two nodes depending on how the CAM model references the CAD model.

After a handle exists for an artifact, when information is required about an artifact, one can

Listing 5.1: Handle syntax from RFC 3651.

```
<Handle> = <NamingAuthority> "/" <LocalName>

<NamingAuthority> = *(<NamingAuthority> ".") <NAsegment>

<NAsegment> = 1*(\%x00-2D / \%x30-3F / \%x41-FF )
              ; any octets that map to UTF-8 encoded
              ; Unicode 2.0 characters except
              ; octets '0x2E' and '0x2F' (which
              ; correspond to the ASCII characters '.',
              ; and '/').

<LocalName> = *(\%x00-FF)
              ; any octets that map to UTF-8 encoded
              ; Unicode 2.0 characters
```

query the handle of an artifact to discover and retrieve its metadata. RFC 3651 provides a set of predefined data types for use in metadata repositories attached to local handle services [22] and this chapter proposes additional data types to address manufacturing contexts. The data types described in Table 5.1 must be included at a minimum for metadata repositories attached to a manufacturing-oriented LHS. In the cases were artifacts are digital objects, the user may retrieve an artifact through the LHS if the user has the appropriate permissions, which invokes the object controller micro-service included in the agent-based adapter to work collaboratively with the local handle services and global identifier (GID) sub-system to resolve the path of the digital object and fetch the complete artifact.

Overall, the user must have the appropriate authentication and authorization to discover and retrieve artifacts. The LHS respects three user access scenarios: (1) objects are not discoverable and retrievable, (2) objects are discoverable and not retrievable, and (3) objects are discoverable and retrievable. The access scenarios respect permissions negotiated by the agent-based adapters and the repositories to which the adapters are attached.

For the work described in the paper, capturing nodes and edges in the local graph database is not automatic. Input describing how to form the nodes and edges was provided to the agent-based adapter. A desired future extension of the LHS is to deploy an inference micro-service that can dynamically track activity in near-real-time and capture autonomously the required information to be curated in the local graph database. Further, several nuances around how industry operates its information technology (IT) networks presents challenges for ensuring effective and persistent identification, addressing, and accessing of artifacts. Section 5.2.2 describes how to overcome the challenges.

```
    ------------------------
   |                        |    4. Result of client request
   | Client with global     |  <-------------------------------.
   |  service information   |                                  |
   |                        |  ----------------------------.   |
    ------------------------     3. Request to responsible |   |
              |   ^                 Local Handle Service   |   |
  1. Client   |   |                                        |   |
  query for   |   |                                        |   |
  naming      |   | 2. Service information                 |   |
  authority   |   |    for "10.1045"                       V   |
  "10.1045"   |   |                     ----------------------
              |   |                    |                      |
              V   |                    | Local Handle Service |
         ---------------               | responsible for the  |
        |               |              | naming authority     |
        | Global Handle |              | "10.1045"            |
        |   Registry    |              |                      |
        |               |               ----------------------
         ---------------
```

Figure 5.3: Handle resolution from global handle registry to local handle service [from [19]].

5.2.2 Persistent Global Identification

The LHS provides a unique, global identifier (GID) system for addressing and searching all artifacts and their inter-relationships in the digital thread. This is a challenging task because the information about the artifacts participating in the digital thread originates from multiple repositories, databases, requirements, system-architecture models, product-structure information (e.g., bill of materials (BOM) and CAD models), and manufacturing plans and data streams. Each type of repository and/or database provides its own identification system that is local to artifacts and relationships managed by such tools. However, when building a digital thread by federating artifacts from multiple repositories, there is a need for a GID system that can be used to address any artifact or inter-relationship throughout the product lifecycle. The Handle System addresses a majority of the tasks for unique, persistent addressing, but the Handle System does not meet all of the manufacturing sector's needs.

The digital thread must work in a global environment of computer firewalls, network security, and multi-layer authentication. The GID for an artifact in the digital thread may not be a single URI as in a generic handle approach. Instead, the GID is often an ordered set of addresses (e.g., URIs or other identifier types) that must be resolved recursively to navigate through multiple layers of namespaces, firewalls, and authentication servers. The digital

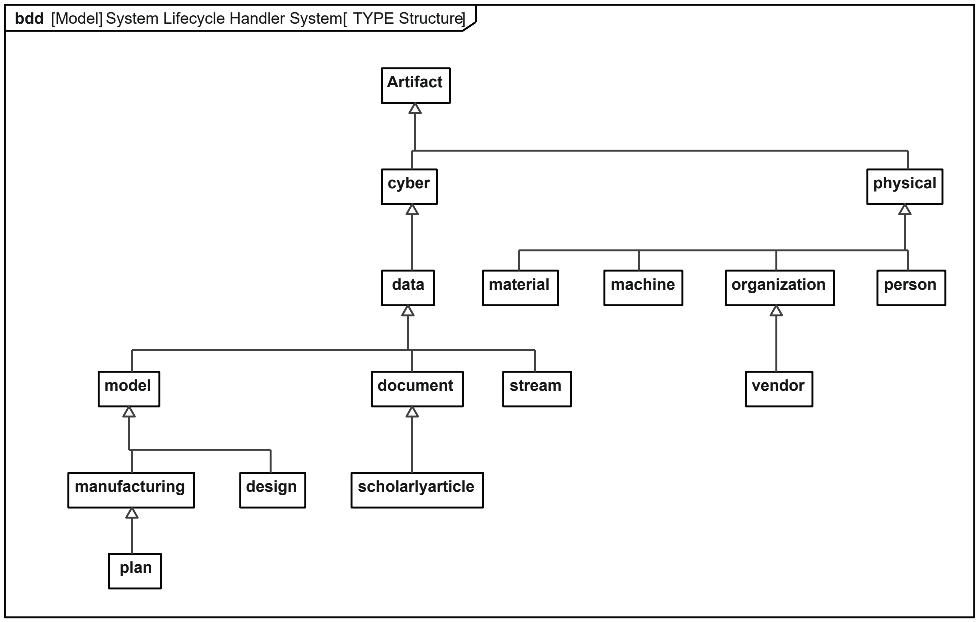


Figure 5.4: Preliminary TYPE structure proposed for describing artifacts referenced in the LHS.

thread includes data originating from multiple sources, such as static files (e.g., spreadsheets, documents), computer models, real-time data streams, and hardware. The identifier for the data may range from cells in a spreadsheet, to unique string-based identifiers for a part in a computer model, to URIs provided by a Representational State Transfer (REST)-based, Hypertext Transfer Protocol (HTTP) service.

Consider the example shown in Figure 5.5 where the design model of a part managed in a PDM system needs to be assigned to a specific machine in the factory that will make this part. For simplicity, assume that the same organization is designing and manufacturing the part. In a globally distributed supply chain, the challenge presented here will be compounded. The abbreviation `A(x,base)` is used to represent the address of an artifact x in the context of the base artifact. The address can be a URI or some form of an identifier that can be resolved. At the highest level, an organization artifact *Org* may have a gateway server available on the internet (world wide web) for all incoming requests, denoted as `A(Org, www)` in the figure. Next, the gateway servers for the various divisions in the organization are generally not reachable directly from the open internet due to firewalls, but reachable from the organization's gateway server. `A(DesignDiv,Org)` and `A(ManufDiv,Org)` are the respective addresses of the design and manufacturing divisions in context of the organization. Similarly, `A(PDM,DesignDiv)` is the address of the PDM server reachable from the design division, and `A(P,PDM)` is the address of the part P in the context of the PDM server. Hence, the GID and address for part P is an ordered set of addresses: `{A(Org,www), A(DesignDiv,Org), A(PDM,DesignDiv), A(P,PDM)}`.

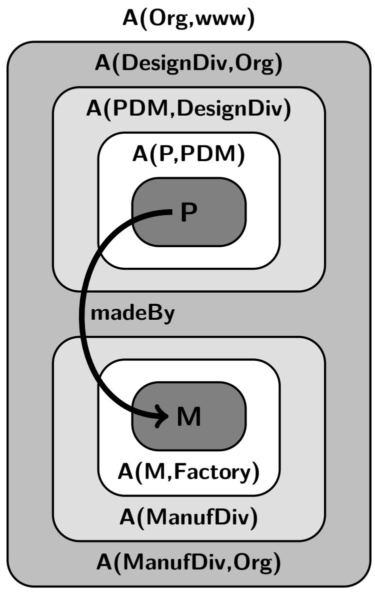


Figure 5.5: Multi-level addresses for locating artifacts across enterprise layers

To reach part P, the LHS is augmented by a resolver system that recursively traverses the chain of addresses, authenticating the request at each base artifact to reach the next artifact. The process for traversing recursively the chain of addresses is shown in Algorithm 1. A similar resolution process must be followed to reach machine M. Once the part P and machine M can be reached uniquely in this manner, a relationship that part P is made by machine M can be established. A collection of these relationships spins the digital thread.

5.3 Case Study

A case study using data from a real design and manufacturing process was used to test and validate the method described in this chapter. The use case is an enclosure box for a payload assembly used in a configurable unmanned aerial vehicle (cUAV). The payload assembly is a subsystem of the overall cUAV system. Figure 5.6 shows the structure of the cUAV in a Systems Modeling Language (SysML) block definition diagram (BDD). The enclosure box is an assembly composed of eight components – four design-build parts and four standard procured parts. This case study focused on only three of the design-build parts: 1) box, 2) internal plate, and 3) cover.

The dataset for the case study comes from the International Council on Systems Engineering (INCOSE) 2015 model-based systems engineering (MBSE) Model Lifecycle Management Workshop [24] and a collaborative project between National Institute of Standards and Technology (NIST) and the Manufacturing Technology Centre (MTC) [25]. Several data types are included in the dataset. The SysML model of the cUAV was retrieved from the results

Algorithm 1: Algorithm for a resolver system that recursively traverses the chain of addresses.

```
Input: A as {GID}
Output: recursive traversal of A
initialization;
if (A = null) then
    return
end
s ← empty stack;
s.push(A);
set s.pos to first address of s;
while s.pos is not end of stack do
    read address at s.pos;
    send request to address at s.pos;
    move s.pos to next address of s;
end
```

of the INCOSE workshop. All other data was retrieved from the NIST-MTC project. Using Solidworks 2016, CAD models captured the digital product definition in accordance with American Society of Mechanical Engineers (ASME) Y14.41-2012 [26]. The manufacturing operations were executed using numerical control (NC) programs in an ISO 6983 [27] compliant format. The manufacturing execution was monitored using an Extensible Markup Language (XML)-based implementation of the MTConnect version 1.3 standard [28]. First article inspection reporting (FAIR) and receiving and incoming inspection (RII) reports were produced using the QIF version 2.1 standard [29].

Figure 5.7 depicts the layered approach for organizing the digital artifacts using in the case study. Three layers were used. The top layer is the Product Concept Level, which contains the high-level stakeholder needs and product requirements. The requirements for the case study were managed in Jama requirements management tool[3]. The middle layer is the Design Variant Level, which contains the digital product definitions and specifications. Four variants of the assembly, four variants of the box, one variant of the plate, and two variants of the cover were available. The CAD models were managed in a GitHub[4] repository and the status of each variant was managed in the Jira issue and project tracking software[5]. The bottom layer is the Part Instant Level, which contains information about the realized product and parts. Twenty instances of each part were fabricated. The MTConnect data, QIF FAIR reports, and QIF RII reports, were managed in a GitHub repository.

Artifacts were linked and managed using the method described in Section 5.2. The imple-

[3] https://www.jamasoftware.com
[4] https://github.com
[5] https://www.atlassian.com/software/jira

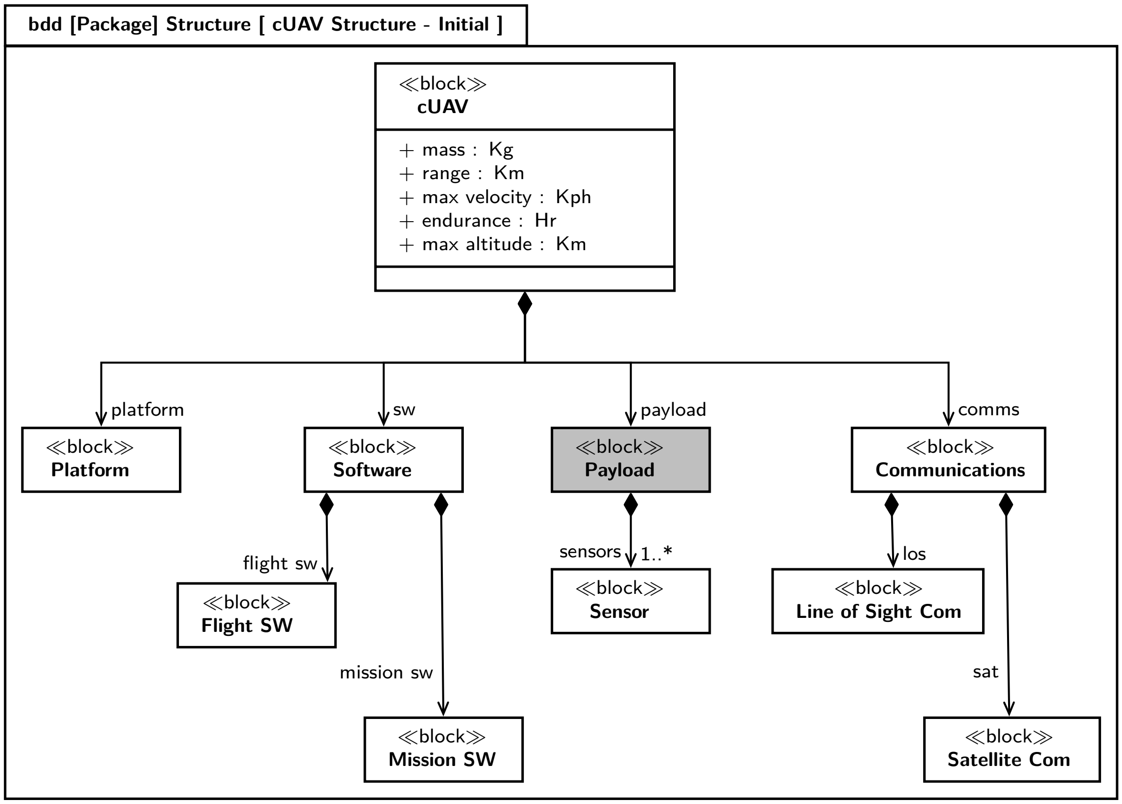


Figure 5.6: A SysML BDD of the overall structure for a cUAV from the INCOSE 2015 MBSE Model Lifecycle Management Workshop [24]. The highlighted "payload" block is the focus of the use case for the case study described in this chapter.

mentation prototype of the architecture and services used several commercially and/or freely available software tools. The Syndeia tool[6] was used to build and manage all the links between the enclosure box assembly artifacts. The graph database was built in Neo4j[7]. Cypher query language [30] was used to query the database.

For this case study, only the enclosure box assembly, box, and cover artifacts were managed in the prototype. Overall, 145 nodes and 436 edges were generated between the assembly, box, and cover. Figure 5.8 presents a chord plot of all the inter-model connections for the assembly, box, and cover. Reviewing a single component in the assembly, Figure 5.9 shows a tree expansion for the connections related to the cover part. There are 44 nodes and 66 edges related to the cover part. When considering a larger product, such as an aircraft or automobile, the number of connections could grow into the thousands, if not millions. The chord plot (Figure 5.8) and tree expansion (Figure 5.9) present evidence to the significant amount of data and links that must be managed, which requires considerable amounts of human resources if managed manually. Searching for and retrieving the right data for a

[6] http://intercax.com/products/syndeia/
[7] https://neo4j.com/product/

particular purpose could take a person hours up to days [31]. Using the method described in this chapter, discovering information, retrieving it, and extracting knowledge took seconds to complete.

Listing 5.2 provides several Cypher queries used for discovering information, retrieving it, and extracting knowledge in this case study. The goal of this case study is to show traceability from multiple viewpoints of the product lifecycle. Using the query on line 2 of Listing 5.2, Figure 5.10 shows all the nearest and next-nearest neighbors to the part instance of the assembly with serial number "D01." Using the query on line 5 of Listing 5.2, Figure 5.11 shows all the tasks connected to the part instance for the box component with serial number "D01." Using the query on line 8 of Listing 5.2, all the design variants of the product concept for the box and the associated CAD files are displayed in Figure 5.12. Using the query on line 11 of Listing 5.2, all part instances of the design variant for "Revision D" of the box are shown in Figure 5.13.

Data traceability can be displayed starting with manufacturing and quality viewpoints too. Using the query on line 14 of Listing 5.2, all the manufacturing and quality files managed and associated with the part instance for the assembly with serial number "D01" are shown in Figure 5.14. Using the query on line 17 of Listing 5.2, Figure 5.15 shows the CAD files managed and associated with the manufacturing data for the box with serial number "5." Using the query on line 20 of Listing 5.2, Figure 5.16 shows all the requirements connected to the part instance of the box with serial number "D01."

Lastly, one part instance of the box was misplaced during shipping and did not go through RII. Using the query on line 23 of Listing 5.2, it can be determined that 20 boxes were fabricated. However, using the query on line 26 of Listing 5.2, it can be seen that only 19 boxes went through RII. The 19 instances of the box can be listed using the query on line 29 of Listing 5.2 and a snippet of the result of the query is shown in Listing 5.3.

The example Cypher queries present evidence of the power of graphs applied to manufacturing contexts. Each node also has a handle associated to it, which provides additional metadata and linking capabilities to enable quickly identifying the type of artifact being referenced. Since the data for the enclosure box resides across several systems in different enterprises, the multi-level addressing method shown in Figure 5.5 was required. An example of the GID for the graph of associated CAD files up through the design variant to the product concept for Revision D of the box component is:

{`20.500.11993/NIST.MTC.CRADA.BOX`,
`20.500.11993/NIST.MTC.CRADA.BOX.SPECIFICATION`,
`20.500.11993/NIST.MTC.CRADA.BOX.REV.D`}.

Listing 5.4 presents a snippet the JSON result for the handle "20.500.11993/ nist.mtc.crada.box.rev.d" from the REST-based API [8].

[8]Manufacturing-related handle metadata can be resolved against the local handle servive API at `https://hdl.mfg.io/api/handles/`.

Listing 5.2: Cypher Query Language entries for the prototyped tested in the case study.

```
// Show all nearest and next-nearest neighbors to Part Instance `NMC_ASSBLY_D01'
MATCH (n1)<-[r1]-(n)-[r]-(s:Block) WHERE s.name=~'NMC_ASSBLY_D01' AND NOT n:Repository AND NOT n:Package AND NOT n1:Package AND NOT n1:Repository RETURN n1,r1,n,r,s

// Show all JIRA Tasks connected to Part Instance `NMC_BOX_D01'
MATCH (m:JIRA_Task)-[r]-(s1)-[r1]-(s:Block) WHERE s.name=~'NMC_Box_D01' RETURN m,r,s1,r1,s

// Show all Design Variants of Product Concept `NIST_MTC_CRADA_BOX' and associated CAD files
MATCH (m:File)-[r1]-(n:Block)-[r:Allocate]-(s:Block) WHERE s.name=~'NIST_MTC_CRADA_BOX' RETURN m,r1,n,r,s

// Show all Part Instances of Design Variant block `NIST_MTC_CRADA_BOX RevD'
MATCH (n:Block)<-[r:Allocate]-(m:Block) WHERE m.name=~'NIST_MTC_CRADA_BOX RevD' RETURN n,r,m

// Show all manufacturing and quality files in GitHub associated with Part Instance `NMC_ASSBLY_D01'
MATCH (m:GitHub_File)<-[t]-(n)<-[r]-(s:Block) WHERE s.name=~'NMC_ASSBLY_D01' RETURN m,t,n,r,s

// Show the CAD files in GitHub associated with manufacturing data `Box-Hurco02-05of20.xml'
MATCH (m:File)-[t1]-(n1)-[t]->(n)-[r]->(s:GitHub_File) WHERE s.name='Box-Hurco02-05of20.xml' RETURN m,t1,n1,t,n,r,s

// Show all Jama requirements connected to Part Instance `NMC_Box_D01'
MATCH (n:Jama_Requirement)-[r2]-(s2)-[r1]-(s1)-[r]-(s:Block) WHERE s.name=~'NMC_Box_D01' AND NOT s1:Repository RETURN n,r2,s2,r1,s1,r,s

// How many instances of Box were fabricated?
MATCH (m:Block)-[t:Allocate]-(n:Block)-[r]-(s:Block) WHERE s.name=~'NIST_MTC_CRADA_BOX' RETURN count(m)

// How many instances of Box were through receiving and incoming inspection?
MATCH (n1:GitHub_File)-[r1:REFERENCE_CONNECTION]-(m:Block)-[t:Allocate]-(s:Block) WHERE n1.name=~'BoxResults_19_samples.QIF' AND s.name=~'NIST_MTC_CRADA_BOX RevD' RETURN count(m)

// List the instances of the box that went through receiving and incoming inspection
MATCH (n1:GitHub_File)-[r1:REFERENCE_CONNECTION]-(m:Block)-[t:Allocate]-(s:Block) WHERE n1.name=~'BoxResults_19_samples.QIF' AND s.name=~'NIST_MTC_CRADA_BOX RevD' RETURN m
```

Listing 5.3: Snippet of text result of a query to list the 19 RII quality files for the box part instance.

```
 "m"
{
 "name": "NMC_Box_D019",
 "gid": "PROJECT-b11f2583-da67-4515-b8d9-1304d22c06a7 |
     _18_5_3_63e021c_1521994265350_944094_15411"
}
{
 "name": "NMC_Box_D018",
 "gid": "PROJECT-b11f2583-da67-4515-b8d9-1304d22c06a7 |
     _18_5_3_63e021c_1521994264434_175058_15408"
}
{
 "name": "NMC_Box_D017",
 "gid": "PROJECT-b11f2583-da67-4515-b8d9-1304d22c06a7 |
     _18_5_3_63e021c_1521994264123_336079_15405"
}
{
 "name": "NMC_Box_D016",
 "gid": "PROJECT-b11f2583-da67-4515-b8d9-1304d22c06a7 |
     _18_5_3_63e021c_1521994263698_60315_15402"
}
.
.
.
```

Listing 5.4: Snippet of JSON result from the REST-based application programming interface (API) for the “20.500.11993/nist.mtc.crada.box.rev.d” handle.


```
{
   "responseCode":1,
   "handle":"20.500.11993/nist.mtc.crada.box.rev.d",
   "values":[
      {
         "index":1,
         "type":"URL",
         "data":{
            "format":"string",
            "value":"https://smstestbed.nist.gov/tdp/mtc/CAD/NIST-MTC-CRADA-Mo...
         },
         "ttl":86400,
         "timestamp":"2018-08-24T19:47:28Z"
      },
      {
         "index":2,
         "type":"TYPE",
         "data":{
            "format":"string",
            "value":"cyber.data.model.design"
         },
         "ttl":86400,
         "timestamp":"2018-08-24T19:47:28Z"
      },
      {
         "index":3,
         "type":"SCHEMA",
         "data":{
            "format":"string",
            "value":"http://schema.org/ProductModel"
         },
         "ttl":86400,
         "timestamp":"2018-08-24T19:47:28Z"
      },
```

This case study shows the magnitude of data that must be managed in manufacturing, even when the dealing with a relatively small product assembly. The case study provides evidence that the method proposed in this chapter potentially overcomes the challenges associated with managing manufacturing-related data [4, 5, 6, 7]. The combination of the graphs and handles associated to each node enables rapid querying, discovering, and retrieving of artifacts based on the users access permissions when proper links are established between related nodes (see Section 5.2.1.

5.4 Discussion

Purpose-built modeling is currently the recommended approach because it enables “expert systems” that supports making decisions in contextual ways related to a specific function and role [32]. Conversely, purpose-built models are not scalable. Data requires context when related to decisions [33]. Data alone is not sufficient for decision making because the decision maker must understand the scope and type of the problem the decision is intended to solve. As the scope of the problem changes, the models must also change. Thus, connecting of heterogeneous information and systems introduces a paradox to the steadfast approach of purpose-driven modeling. A trade-off of how purpose-built to make a model versus how scalable (i.e., generalized) to make a model must be considered. This requires integrating domains in multiple directions while providing scalable contextual models. Overcoming these challenges is not easy, but a standards-based linked-data approach, using the digital thread, provides the best opportunity for maximizing the successful deployment of smart manufacturing.

5.4.1 Contextualizing Graph-Based Viewpoints

Using graphs to link data across the product lifecycle provides the value of being able to quickly extract domain-specific viewpoints. For instance, consider the graphs shown in Figure 5.17. Figure 5.17(a) provides a product lifecycle viewpoint for the box component from the case study in Section 5.3. While Figure 5.17(b) and (c) present an emphasized subgraph for a manufacturing-specific viewpoint and a materials-specific viewpoint, respectively.

Enabling the viewpoint-identification capability provides efficient and effective segmentation of the massive datasets generated by enterprises in the manufacturing sector. Without this approach, industry spends considerable amounts of time searching for data related to its products. With this approach , domain experts can quickly find information from across the enterprise that relates to their needs and they can quickly move to gathering actionable intelligence.

5.4.2 Knowledge Generation

Traceability, impact analysis, and continuous validation and integration of the digital thread are important aspects of configuration management. The greatest impact of the digital thread is in the continuous analyses that can be performed. In the simplest form, basic traceability exists where one can traverse the digital thread using the intra- and inter-model connections, starting with any artifact. However,a greater capability is to use graph-pattern matching and graph traversals to assess the upstream and downstream impact of changes in any artifact [15]. For example, computing the downstream impact of changes in a requirement, or querying upstream requirements and analyses done on a part when it fails during operation.

Further, Feng et al. [10] developed a method for managing knowledge in the context of smart manufacturing. The authors provided three contributions: (1) context for data, information, understanding, and autonomy in knowledge generation, (2) knowledge constructs decomposed into basic, composable units, and (3) a reference application to smart manufacturing. However, Feng et al. found further advances in knowledge-base architectures is required to better enable integration of information across a product lifecycle [10]. This chapter has shown through the case study in Section 5.3 that distributed and/or federated information can be effectively linked across several enterprises and information can quickly be curated, discovered, and retrieved. Feng et al. defined necessary knowledge contructs for the product lifecycle, while the work described here provides the necessary information structuring, object representation, and communication mechanisms. Together, the two bodies of work provide a viable solution to industry for enabling smart manufacturing digital threads.

5.4.3 Further Research

Dynamically Generating Connections. The LHS must provide capabilities to generate and register artifacts in the digital thread and to link them using connections. This includes, for example, generating design models from requirements (e.g., design synthesis), or generating simulation models and manufacturing process plans from design models, or registering new machines and machine configurations on a factory floor. Further, the LHS should enable automated generation of connections between artifacts when one is generated from the other. For instance, connections between design and manufacturing models are automatically generated when manufacturing models are generated from design models. This would overcome the manual creation of connections between artifacts that is laborious. Further research is required to enable the autonomous linking capabilities. Specifically, in near-real-time, how are all the links across enterprises tracked? Or how can inference systems be used to facilitate tracking of links?

Frequently Asked Queries. This chapter presented example queries through the case study described in Section 5.3. However, a complete and concise set of recommended "Frequently Asked Queries" must be research furthered to provide industry with a reference library of graph-based queries that can be deployed to answer key questions across the product lifecycle. Each role in the product lifecycle has typical questions he/she asks about a product while executing tasks in the context of his/her domain expertise. Combining a library of common queries with the methods described in this chapter could significantly reduce the effort of human capital in making decisions by leveraging the capabilities of generating contextual graph-based viewpoints and quickly extracting actionable intelligence through knowledge generation.

5.5 Conclusion

This chapter provides a method for using graphs to link data across the product lifecycle for enabling smart manufacturing digital threads. It allows the possibility of quickly locating artifacts across distributed and/or federated enterprises without making any presumptions about the objects of the artifacts or their locations. The major contributions of this work are a standards-based, linked-data approach, providing seamless traceability across the product lifecycle, enabling high-quality manufacturing contextualization of information, and supporting enterprise-wide knowledge reuse.

The method presented here leverages several established and trusted approaches and technologies. The first is the Handle System, which is the backbone to the widely popular Distributed Object Identifier (DOI) system that persistently identifies media and publication objects – for instance, the publisher of this chapter provided a DOI for pointing universally to the paper. Second, the foundations of graph-theory are leveraged, which provides the ability to quickly make, track, and query connections between artifacts in support of contextually generating knowledge about the product lifecycle. Last, generally accepted linked-data approaches are extended to manufacturing contexts and provided additional capabilities to overcome architectural and IP-related challenges that are specific to the manufacturing section.

The next steps in this work is to enhance the reliability of data available by introducing more rigor in how links are stored, configured, and where the links are stored. Further, a more comprehensive metadata schema is in development. This includes leveraging work to extract the minimum information requirements to complete one loop of the product lifecycle [7, 9, 34]. Making these enhancements puts the LHS in a good position to deliver significant impact through enabling cost-effective deployment of digital threads.

Chapter Bibliography

[1] MAPI Foundation. Facts about modern manufacturing, 2015. URL https://mapifoundation.org/manufacturing-facts/2015/8/9/facts-about-modern-manufacturing.

[2] Thomas Hedberg Jr, Joshua Lubell, Lyle Fischer, Larry Maggiano, and Allison Barnard Feeney. Testing the digital thread in support of model-based manufacturing and inspection. *Journal of Computing and Information Science in Engineering*, 16(2):021001, mar 2016. ISSN 1530-9827. doi: 10.1115/1.4032697.

[3] Gary Anderson. The economic impact of technology infrastructure for smart manufacturing. Technical report, National Institute of Standards and Technology, Gaithersburg MD, 2016. URL http://nvlpubs.nist.gov/nistpubs/eab/NIST.EAB.4.pdf.

[4] Biren Prasad, Roger S. Morenc, and Ravi M. Rangan. Information management for concurrent engineering: Research issues. *Concurrent Engineering*, 1(1):3–20, mar 1993. ISSN 1063-293X. doi: 10.1177/1063293X9300100102. URL http://journals.sagepub.com/doi/10.1177/1063293X9300100102.

[5] Ravi M. Rangan, Steve M. Rohde, Russell Peak, Bipin Chadha, and Plamen Bliznakov. Streamlining product lifecycle processes: A survey of product lifecycle management implementations, directions, and challenges. *Journal of Computing and Information Science in Engineering*, 5(3):227, sep 2005. doi: 10.1115/1.2031270.

[6] Patrick Waurzyniak. Manufacturing factory data. *Manufacturing Engineering*, pages 71–82, jul 2012. URL http://www.sme.org/uploadedFiles/Publications/ME{_}Magazine/2012/July{_}2012/July2012f3Software.pdf.

[7] Thomas D Hedberg Jr, Nathan W Hartman, Phil Rosche, and Kevin Fischer. Identified research directions for using manufacturing knowledge earlier in the product life cycle. *International Journal of Production Research*, 55(3):819–827, 2017. doi: 10.1080/00207543.2016.1213453.

[8] Asa Trainer, Thomas Hedberg Jr, Allison Barnard Feeney, Kevin Fischer, and Phil Rosche. Gaps analysis of integrating product design, manufacturing, and quality data in the supply chain using model-based definition. In *ASME 2016 11th International Manufacturing Science and Engineering Conference – Volume 2: Materials; Biomanufacturing; Properties, Applications and Systems; Sustainable Manufacturing*, volume 2, page V002T05A003. ASME, jun 2016. doi: 10.1115/MSEC2016-8792.

[9] Shawn P. Ruemler, Kyle E. Zimmerman, Nathan W. Hartman, Thomas Hedberg Jr, and Allison Barnard Feeney. Promoting model-based definition to establish a complete product definition. *Journal of Manufacturing Science and Engineering*, 139(5):051008, nov 2016. ISSN 1087-1357. doi: 10.1115/1.4034625.

[10] Shaw C. Feng, William Z. Bernstein, Thomas Hedberg Jr, and Allison Barnard Feeney. Toward knowledge management for smart manufacturing. *Journal of Computing and Information Science in Engineering*, 17(3):031016, jul 2017. doi: 10.1115/1.4037178.

[11] William Regli, Jarek Rossignac, Vadim Shapiro, and Vijay Srinivasan. The new frontiers in computational modeling of material structures. *Computer-Aided Design*, 77:73–85, 2016. doi: 10.1016/j.cad.2016.03.002.

[12] Timothy Sprock, Anike Murrenhoff, and Leon F. McGinnis. A hierarchical approach to warehouse design. *International Journal of Production Research*, 55(21):6331–6343, nov 2017. ISSN 0020-7543. doi: 10.1080/00207543.2016.1241447.

[13] Timothy Sprock and Conrad Bock. Incorporating abstraction methods into system-analysis integration methodology for discrete event logistics systems. In *2017 Winter Simulation Conference (WSC)*, pages 966–976, Las Vegas, NV, dec 2017. IEEE. ISBN 978-1-5386-3428-8. doi: 10.1109/WSC.2017.8247847.

[14] Manas Bajaj, Dirk Zwemer, Rose Yntema, Alex Phung, Amit Kumar, Anshu Dwivedi, and Manoj Waikar. MBSE++ – foundations for extended model-based systems engineering across system lifecycle. In *INCOSE International Symposium*, volume 26, pages 2429–2445, Edinburgh, Scotland, UK, jul 2016. Wiley-Blackwell. doi: 10.1002/j.2334-5837.2016.00304.x.

[15] Manas Bajaj, Jonathan Backhaus, Tim Walden, Manoj Waikar, Dirk Zwemer, Chris Schreiber, Ghassan Issa, and Lockheed Martin. Graph-based digital blueprint for model based engineering of complex systems. In *INCOSE International Symposium*, volume 27, pages 151–169, Adelaide, Australia, jul 2017. Wiley-Blackwell. doi: 10.1002/j.2334-5837.2017.00351.x.

[16] Manas Bajaj and Thomas Hedberg Jr. System lifecycle handler – spinning a digital thread for manufacturing. In *28th Annual INCOSE International Symposium*, pages 1636–1650, Washington, DC, 2018. Wiley. doi: 10.1002/j.2334-5837.2018.00573.x.

[17] Robert Kahn and Robert Wilensky. A framework for distributed digital object services. *International Journal on Digital Libraries*, 6(2):115–123, apr 2006. doi: 10.1007/s00799-005-0128-x.

[18] Thomas D Hedberg Jr, Allison Barnard Feeney, Moneer M Helu, and Jaime A Camelio. Toward a lifecycle information framework and technology in manufacturing. *Journal of Computing and Information Science in Engineering*, 17(2):021010, feb 2017. ISSN 1530-9827. doi: 10.1115/1.4034132.

[19] Sam. Sun, Larry. Lannom, and Brian. Boesch. Handle system overview. Technical report, The Internet Society, Network Working Group, nov 2003. URL https://www.rfc-editor.org/info/rfc3650.

[20] International Standards Organization. Information and documentation – digital object identifier system, 2012.

[21] International DOI Foundation. Key facts on digital object identifier system, 2018. URL http://www.doi.org/factsheets/DOIKeyFacts.html.

[22] Sam. Sun, Sean Reilly, and Laurence Lannom. Handle system namespace and service definition. Technical report, The Internet Society, Network Working Group, nov 2003. URL https://www.rfc-editor.org/info/rfc3651.

[23] Sam. Sun, Sean. Reilly, Larry. Lannom, and Jason. Petrone. Handle system protocol (ver 2.1) specification. Technical report, The Internet Society, Network Working Group, nov 2003. URL https://www.rfc-editor.org/info/rfc3652.

[24] Amit Fischer and Gavin Arthurs. INCOSE 2015 MBSE workshop breakout session, 2015. URL http://www.omgwiki.org/MBSE/doku.php?id=mbse:incose{_}mbse{_}iw{_}2015:breakout{_}out{_}session{_}model{_}lifecylce{_}mgmt.

[25] Thomas Hedberg Jr, Michael Sharp, Toby Maw, Mostafizur Rahman, Swati Jadhav, James Whicker, Allison Barnard Feeney, and Moneer Helu. A three component assembly with design, manufacturing, and inspection data from a collaboration between the National Institute of Standards and Technology and the Manufacturing Technology Centre. *Journal of Research of National Institute of Standards and Technology*, 123 (123xxx), 2018. doi: 10.6028/jres.123.XXX.

[26] American Society of Mechanical Engineers. Digital Product Definition Data Practices, 2012.

[27] International Standards Organization. Automation systems and integration – Numerical control of machines – Program format and definitions of address words – Part 1: Data format for positioning, line motion and contouring control systems, 2009.

[28] MTConnect Institute. MTConnect Standard, 2014. URL http://www.mtconnect.org/media/39542/mtc_part_1_overview_v1.3.pdf.

[29] Dimensional Metrology Standards Consortium. Part 1: Overview and Fundamental Principles in Quality Information Framework (QIF) – An Integrated Model for Manufacturing Quality Information, 2014. URL http://qifstandards.org/.

[30] Nadime Francis, Andrés Taylor, Alastair Green, Paolo Guagliardo, Leonid Libkin, Tobias Lindaaker, Victor Marsault, Stefan Plantikow, Mats Rydberg, and Petra Selmer. Cypher: An Evolving Query Language for Property Graphs. In *Proceedings of the 2018 International Conference on Management of Data - SIGMOD '18*, pages 1433–1445, New York, New York, USA, 2018. ACM Press. ISBN 9781450347037. doi: 10.1145/3183713.3190657. URL http://dl.acm.org/citation.cfm?doid=3183713.3190657.

[31] Sebastian Adolphy, Hendrik Grosser, Lucas Kirsch, and Rainer Stark. Method for automated structuring of product data and its applications. In *Procedia CIRP*, volume 38, pages 153–158. Elsevier, jan 2015. doi: 10.1016/j.procir.2015.07.063.

[32] Thomas D Hedberg Jr, Moneer M Helu, and Timothy Sprock. A standards and technology roadmap for scalable distributed manufacturing systems. In *Proceedings of the 2018 Manufacturing Science and Engineering Conference*, College Station, Texas, 2018. American Society of Mechanical Engineers (ASME). doi: 10.1115/MSEC2018-6550.

[33] Dennis Brandl. Drowning in data, starved for information. *Control Engineering*, 2013. URL http://www.controleng.com/search/search-single-display/drowning-in-data-starved-for-information/a2d369dd2f82d2cd5788e5b02e1e91ef.html.

[34] Alexander McDermott Miller, Nathan Hartman, Thomas Hedberg Jr, Allison Barnard Feeney, and Jesse Zahner. Towards identifying the elements of a minimum information model for use in a model-based definition. In *ASME 2017 12th International Manufacturing Science and Engineering Conference collocated with the JSME/ASME 2017 6th International Conference on Materials and Processing – Volume 3: Manufacturing Equipment and Systems*, volume 3, page V003T04A017. ASME, jun 2017. doi: 10.1115/MSEC2017-2979.

Table 5.1: The schema for the metadata repositories attached to manufacturing-oriented local handle systems in the LHS system.

Data Type	Index	Requirement	Description
URI	1	Required	The URI type provides a Uniform Resource Identifier (URI) that is passed to a general-purpose name service for accessing the artifact referenced by a handle.
EMAIL	2	Optional	The EMAIL type provides a UTF8-encoded email addresses for a handle that points to a person.
TYPE	3	Optional	The TYPE type provides the type of artifact that is referenced by a handle. The TYPE notation is based on the proposed structure presented in Figure 5.4.
SCHEMA	4	Optional	The SCHEMA type provides the schema used to provide the data provided by the ATTRIBUTE type. The SCHEMA type is required when the ATTRIBUTE type is included for a handle.
DATE_CREATE	5	Optional	The DATE_CREATE type captures the timestamp for when the artifact referenced by the handle was originally created.
ATTRIBUTE	6	Optional	The ATTRIBUTE type provides informative data about the artifact in JavaScript Object Notation (JSON) form according to the schema provided by the SCHEMA type.
ATTRIBUTE64	7	Optional	The ATTRIBUTE64 type is a base64 encoding of the data provided by the ATTRIBUTE type. This data type is intended to enable automation by providing computer-interpretable data.
EFFECTIVITY	10	Optional	The EFFECTIVITY type provides the date effectivity or serial effectivity for the artifact.
HS_ADMIN	100	Required	The HS_ADMIN as defined by RFC 3651 [22].
HS_PUBKEY	300	Required	The HS_PUBKEY type provides encoded information describing a public key for authenticating entities in the handle system.
HS_SIGNATURE	400	Optional	The HS_SIGNATURE type provides the digital signature of an entity that vouches for the metadata included for a handle.

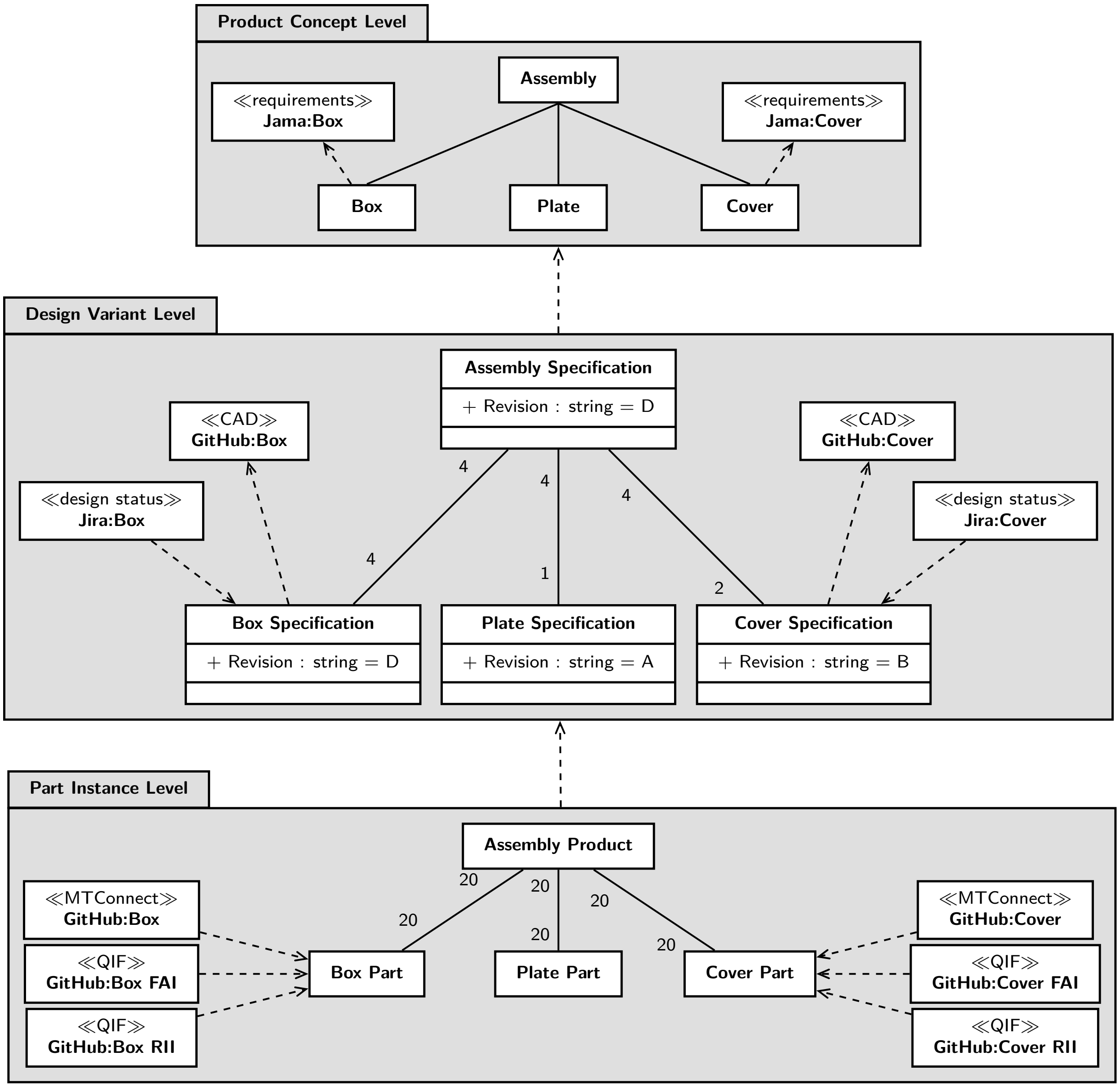


Figure 5.7: An overview of the layered approach for organizing the digital artifacts of the enclosure box for an payload assembly used in the cUAV use case. Data came from the INCOSE 2015 MBSE Model Lifecycle Management Workshop [24] and a collaborative project between NIST and the MTC [25]. FAI = first article inspection, RII = receiving and incoming inspection

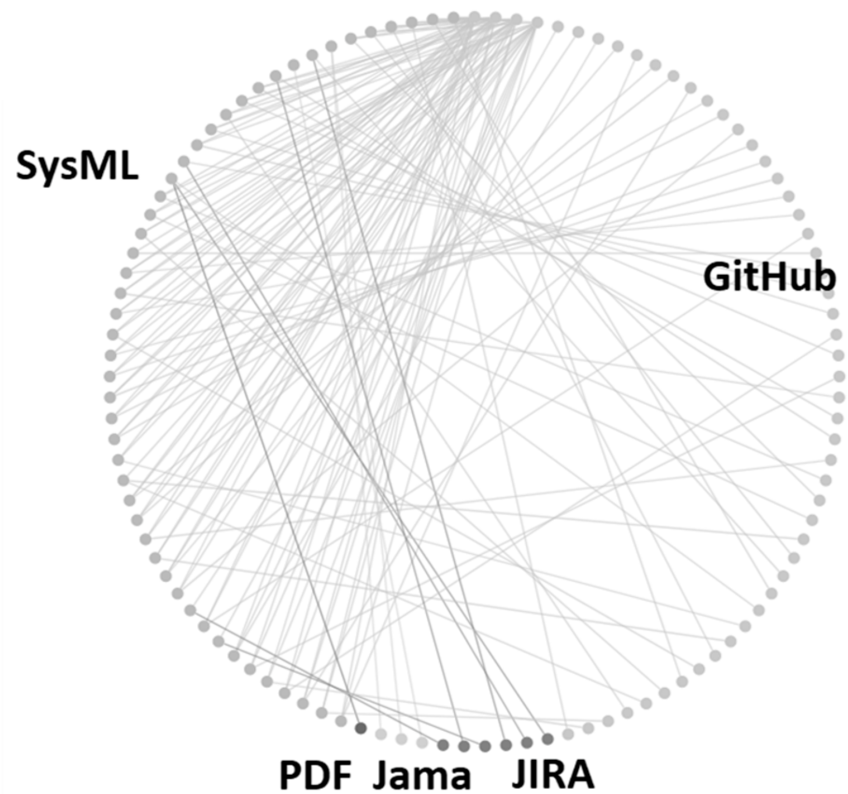


Figure 5.8: Chord plot showing all inter-model connections between the assembly, box, and cover.

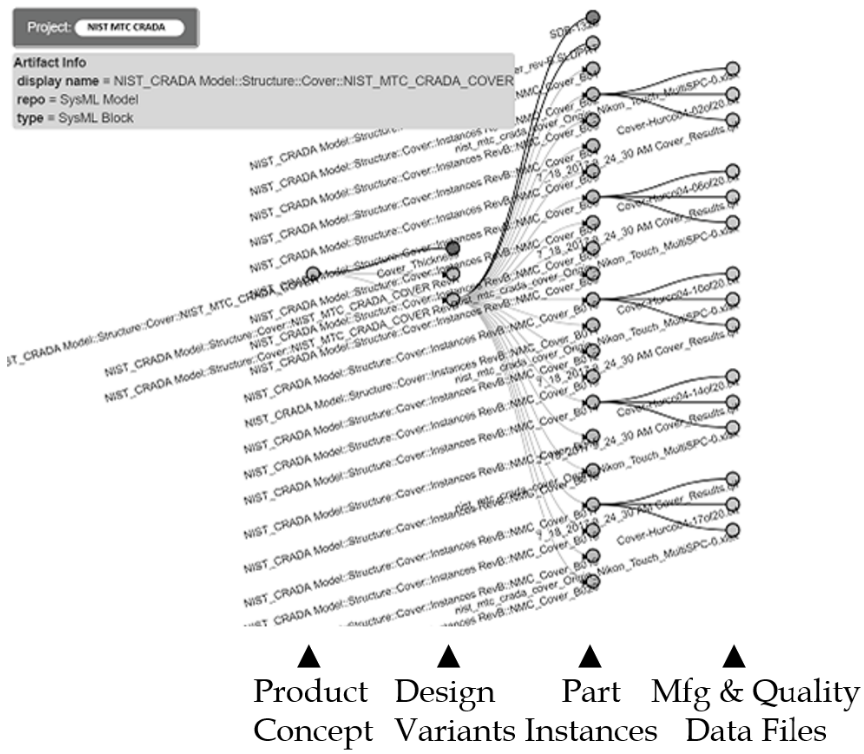


Figure 5.9: Tree expansion from the Product Concept Level of the cover. The tree shows the links between the Product Concept Level through the Design Variant and Part Instance Levels to the individual manufacturing and quality files.

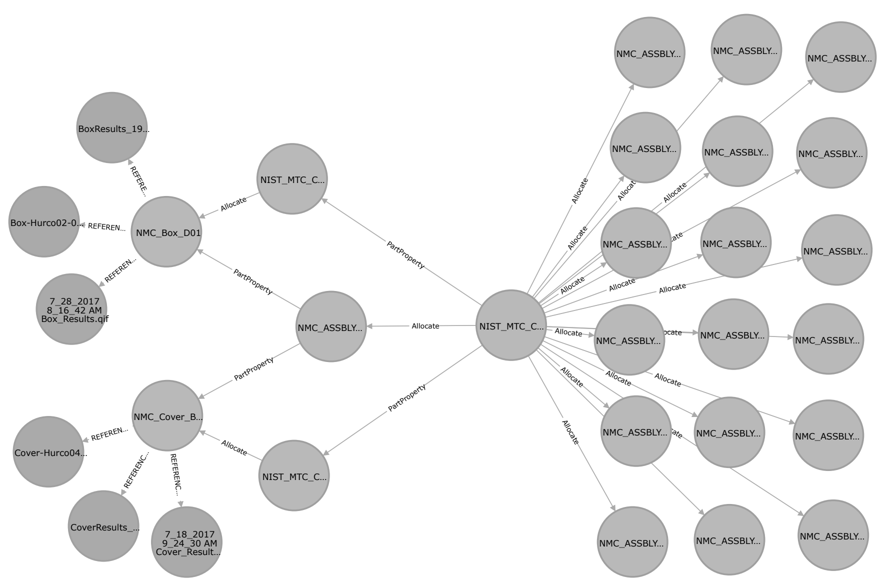


Figure 5.10: Sub-graph showing all nearest and next-nearest neighbors to Part Instance ‘NMC_ASSBLY_D01.’

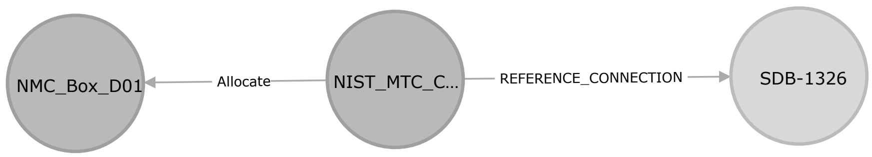


Figure 5.11: Sub-graph showing all JIRA Tasks connected to Part Instance ‘NMC_BOX_-D01.’

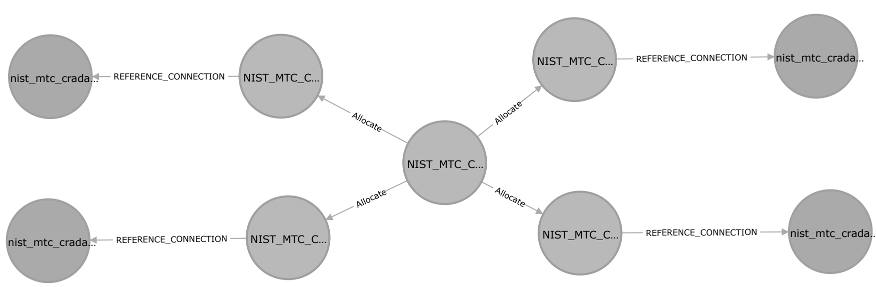


Figure 5.12: Sub-graph showing all Design Variants of Product Concept ‘NIST_MTC_-CRADA_BOX’ and associated CAD files.

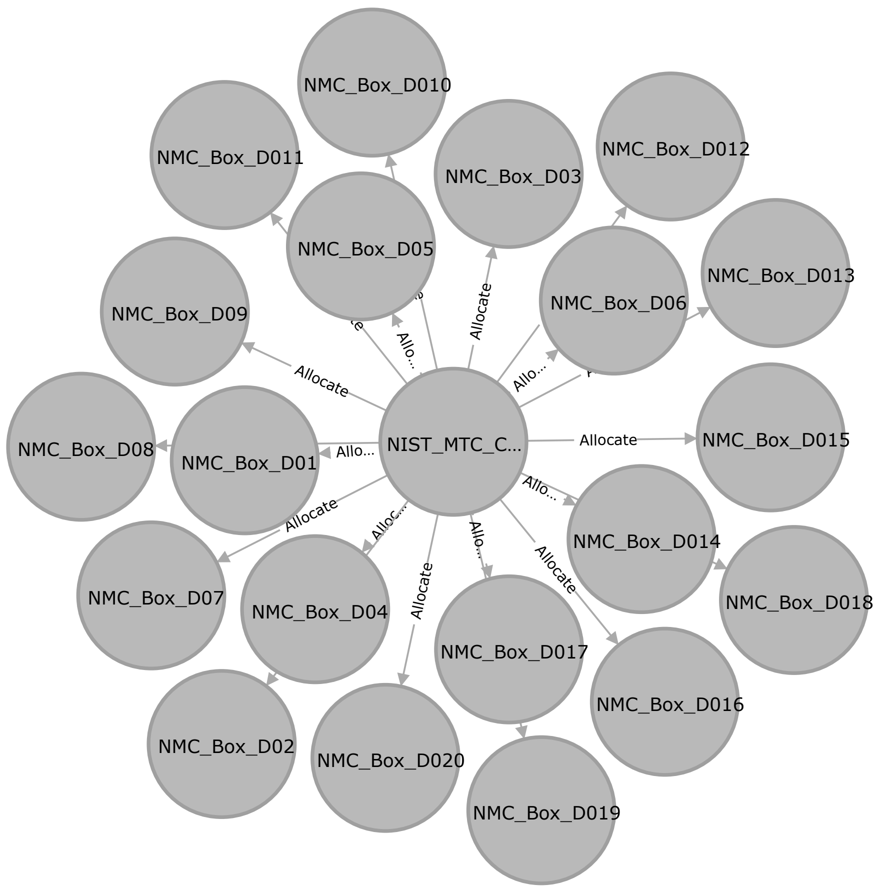


Figure 5.13: Sub-graph showing all Part Instances of Design Variant block 'NIST_MTC_-CRADA_BOX RevD.'

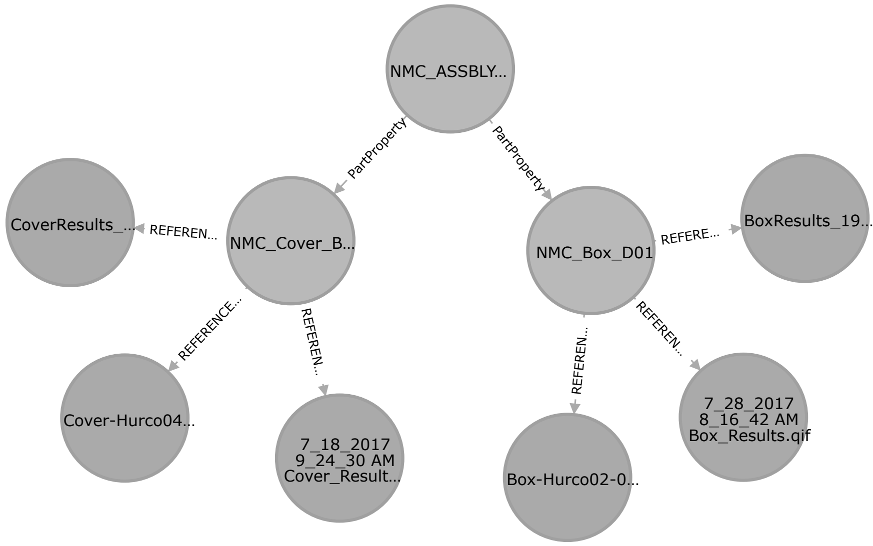


Figure 5.14: Sub-graph showing all all manufacturing and quality files in GitHub associated with Part Instance 'NMC__ASSBLY__D01.'

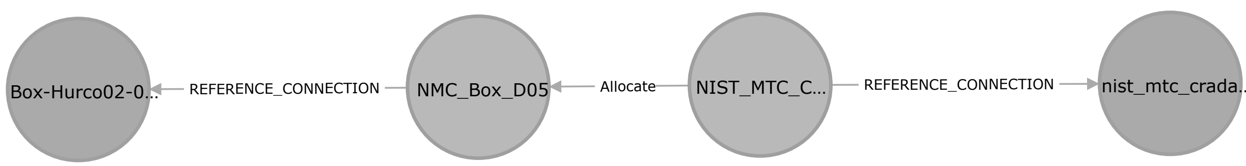


Figure 5.15: Sub-graph showing the CAD files managed and associated with manufacturing data 'Box-Hurco02-05of20.xml.'

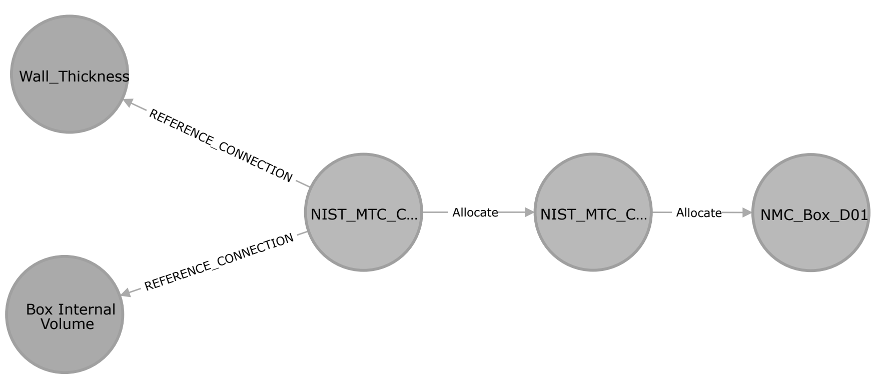


Figure 5.16: Sub-graph showing all all Jama requirements connected to Part Instance 'NMC__Box__D01.'

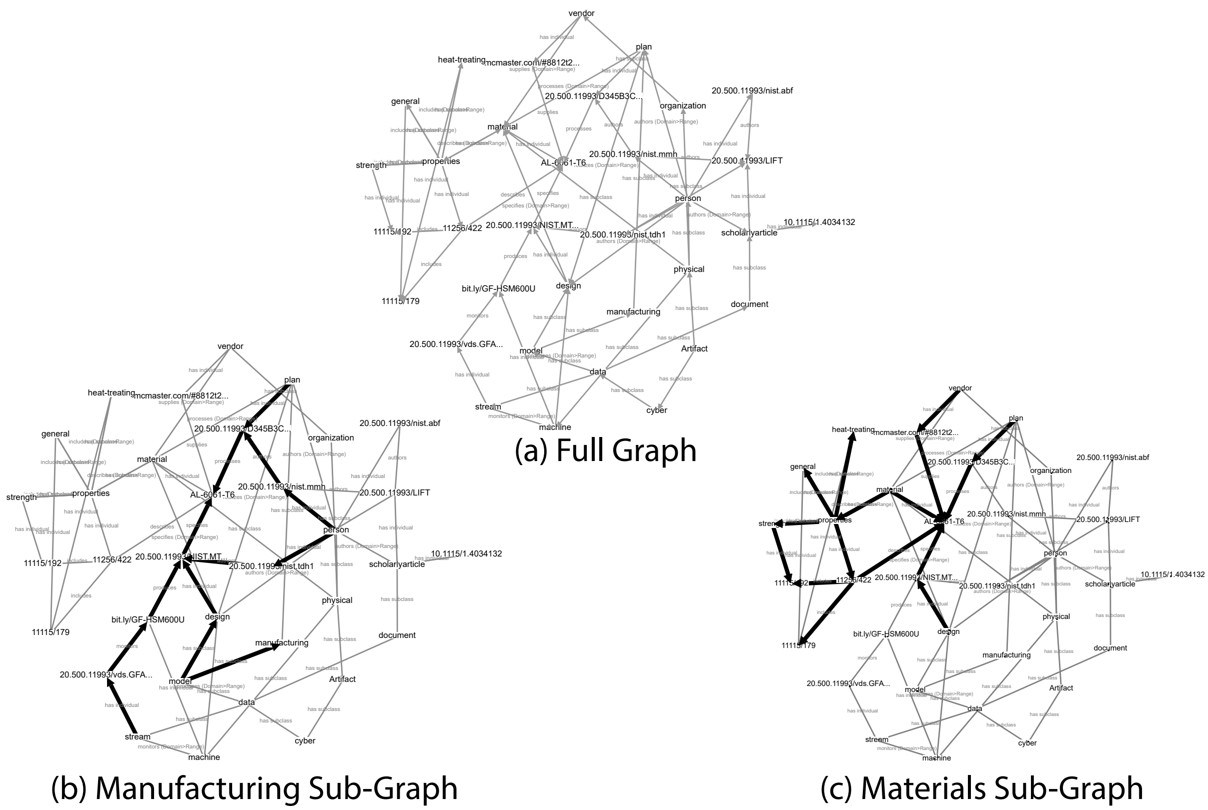


Figure 5.17: Contextual Graph-Based Viewpoints for a full product lifecycle viewpoint, a manufacturing viewpoint, and a materials viewpoint.

Chapter 6

Data Certification and Traceability[1]

6.1 Introduction

Information technology advances such as big data, service-oriented architectures, and networking have triggered a digital revolution [1] that holds promise of reduced costs, improved productivity, and higher quality. Modern manufacturing enterprises are both more globally distributed and more digital than ever before, resulting in increasingly complex manufacturing system networks [2, 3]. Manufacturers are under mounting pressure to perform digital manufacturing more efficiently and effectively within these distributed manufacturing systems. Moreover, engineers are being pushed by industry and business demands to use more manufacturing information and knowledge in their design decisions [4]. To do so, industry is changing how product definitions are communicated – from paper to models.

Those leading the efforts to transition communication methods for manufacturing complex products coined the term "digital thread" to convey the data flows between engineering, manufacturing, business processes, and across supply chains [5]. With the advent of new manufacturing-data standards [6] and more powerful engineering software, it is now possible to perform all engineering functions using a model-based definition (MBD) [4]. A MBD is a three-dimensional (3D) digital product model that defines the requirements and specifications of the product. A model-based enterprise (MBE) approach uses these models, rather than documents, as the data source for all engineering activities throughout the product lifecycle. The core MBE tenets are that models are used to drive all aspects of the product lifecycle and that data is created once and reused by all downstream data consumers.

This transition to a MBE has introduced new requirements on data usage across the product lifecycle. The need for automated methods to collect, transmit, analyze, and act on the most appropriate data is gaining attention in the literature [7, 8, 9, 10]. Research in model-based-data interoperability between design activities (e.g., product and assembly design) and manufacturing activities (e.g., fabrication, assembly, and quality assurance) is also gaining momentum [11]. However, more effort is needed in the area related to trustworthiness to support authentication, authorization, and traceability of product data. Product data must

[1]This chapter was published as an article with the citation: Thomas D. Hedberg Jr, Sylvere Krima, and Jaime A. Camelio. Embedding x.509 digital certicates in three-dimensional models for authentication, authorization, and traceability of product data. *Journal of Computing and Information Science in Engineering*, 17 (1):011008-011008-11, 2016. ISSN 1530-9827. doi: 10.1115/1.4034131.

be guaranteed by an authority to a predefined level of data quality and trustworthiness if that information is to be used throughout the product lifecycle. That is, the user must be able to know who did what to whom and when it was done.

A method and technology to support authentication, authorization, and traceability of product data was developed during the research. This technology enables trust throughout the product lifecycle. Requirements of trustworthiness are not defined, because that work is happening in other places[2]. The aim is to supplement the requirements work by providing the infrastructure for transmitting the information (e.g., provenance, metadata) required to enable trustworthiness in the product lifecycle.

Our methodology and technology follows recommended practices from Semantic Web [12] concepts using the X.509 standard [13]. This standard enables us to embed digital certificates with authentication, authorization, and traceability meta-data into 3D models. An open-source digital manufacturing certificate (DMC) toolkit[3] was developed that provides an application programming interface (API) and user interface to embed digital certificates into four standards-based 3D-model formats. While X.509 has been adopted heavily by the cyber-security domain, the work here is not trying to provide security methods. The goal is to provide a mechanism for a data user to know what the data is (i.e., the authentication), how the data can be used (i.e., the authorization), and what has happen to the data throughout the product lifecycle (i.e., the traceability).

This chapter describes the use of X.509 certificates and proposes a solution for embedding X.509 digital certificates in 3D models for authentication, authorization, and traceability of product data. This paper also describes the application of this technology to an Aerospace part. Finally, the paper draws conclusions, provides recommendations, and details our next steps for further research into using X.509 certificates in product lifecycle management (PLM) workflows to enable trustworthiness throughout the product lifecycle.

6.2 Implementation Description

Digital certification could help to control and improve the product-data quality (PDQ) – the digital backbone of smart manufacturing – throughout the product lifecycle. Smart manufacturing is a recent concept whose requirements and possibilities have not all been explored and standard information frameworks for product data do not currently cover all of them. The Public Key Infrastructure (X.509-PKI) infrastructure was chosen because it is widely adopted across several industries and it is efficient to implement. The method is for digitally signing the data – not encrypting the data. This approach was taken because the methods

[2]The National Institute of Standards and Technology (NIST) Cyber-Physical Systems Public Working Group is working on trustworthiness requirements and frameworks. For more information about the Cyber-Physical Systems Public Working Group goto: `https://pages.nist.gov/cpspwg/`

[3]The toolkit is available at: `https://github.com/usnistgov/DT4SM`

are not trying to provide security mechanisms. There are various activities in industry that are focused on the topic of cyber-security [14] – I do not want to duplicate those efforts. In addition, encryption only provides security at the edge of the communication. Once the data is decrypted for usage, there is limited-to-no way to control how that data usage continues on in the lifecycle. Recall, our goal was to provide a vehicle for transmitting the required information to support the process of trustworthiness in the product lifecycle. Digital signatures conforming to X.509-PKI provide a method for including traceability information in a sustainable way. If a user modifies the data, either intentionally or unintentionally, the digital certificate would become invalid. This provides a level of control and data guarantees that ensure the data user knows who did what to the data and when it was done.

Traceability may be classified across three categories: i) internal/external, ii) forward/backwards, iii) active/passive. Cheng and Simmons [15] defines the first category as both internal and external traceability. Internal traceability is the traceability inside the factory and the product system. External traceability follows the product into its relationships with customers, maintainers, and service providers. The next category comes from Jansen-Vullers et al. [16], where traceability is classified into backward and forward traceability. Backward traceability records information and data on the past history of the product. Forward traceability explains what will happen to a certain product, in terms of operations and processes – this information is written before performing any operation. The last category is active and passive traceability. Active traceability is considered to be on-line and synchronous, which implies the data may be "phoning home" to a central server. Passive traceability may be on-line or off-line and is typically asynchronous. Our method supports all combinations of the traceability categories. The only requirement to using our method for traceability is the availability to validate the attached digital certificates.

The remainder of Section 6.2 describes and discusses the solution. First, is to define the signature block for capturing the digital signature using a digital certificate. Next, a short use case is presented to provide context to implementing our solution. Then, a description of the proposal for extending both the ISO 10303-21 (STEP Part 21) and Quality Information Framework (QIF) standards to support our solution. Lastly, a provided aerospace example to demonstrate the usage of the authentication, authorization, and traceability information.

6.2.1 Signature Block

The DMC toolkit is designed specifically to provide certification – digital signature using software and hardware certificates, and verification – of manufacturing-related data. It enables generating, embedding, and verifying signatures in a manufacturing-related file.

ISO 10303 defines different serialization mechanisms to encode product data in ASCII files, such as ISO 10303-21 [17]. The DMC toolkit follows the work from the ISO Technical Committee 184 / Subcommittee 4 on the ISO 10303-21 3rd edition. This draft recommends to embed the signature, following the PKCS#7 format, at the end the data file, in a signature data block. A signature data block opens with a `SIGNATURE;` tag and closes with an `ENDSEC;` tag. A valid signature block example is shown in Listing 6.1.

Listing 6.1: Signature block for STEP (ISO 10303-21 ed. 3)

```
SIGNATURE;
-----BEGIN PKCS7-----
signature in pkcs7 format
-----END PKCS7-----
ENDSEC;
```

ISO 6983 does not officially provide guidance on how to embed signatures in its data file. The DMC Toolkit implementation follows similar guidelines to ISO 10303-21 3rd edition. In the current version, digital signatures can be found at the end of the file as comments between the `(-----BEGIN PKCS7-----)` and `(-----END PKCS7-----)` tags. A valid signature block would look like the example shown in Listing 6.2.

Listing 6.2: Signature block for ISO 6983

```
(-----BEGIN PKCS7-----)
signature in pkcs7 format
(-----END PKCS7-----)
```

The DMC toolkit signs PDF [18] with embedded PRC [19] files using the iText PDF library[4]. This library follows the official ISO standard, providing interoperable signatures that can be read by any software component compatible with ISO 32000 [18].

The Quality Information Framework uses Extensible Markup Language (XML) to represent its data. The W3C has a standard that describes digital-signature representation for XML data (XMLDsig)[5]. The DMC toolkit follows the XMLDsig specification to encode QIF signatures. QIF signatures are the last XML nodes of the root node of the QIF document. A valid signature example is shown in Listing 6.3.

Listing 6.3: Signature example in a QIF document

```
<?xml version="1.0" encoding="utf-8"?>
<QIFDocument ...>
...
<Signature xmlns="http://www.w3.org/2000/09/xmldsig\#">
<SignedInfo>
<CanonicalizationMethod
    Algorithm="http://www.w3.org/TR/2001/REC-xml-c14n-20010315" />
<SignatureMethod Algorithm="http://www.w3.org/2000/09/xmldsig\#rsa-sha1" />
<Reference URI="">
<Transforms>
<Transform Algorithm="http://www.w3.org/2000/09/xmldsig\#enveloped-signature" />
</Transforms><DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig\#sha1"
    />
```

[4] http://sourceforge.net/projects/itext/

[5] https://www.w3.org/TR/xmldsig-core/

```
<DigestValue>...</DigestValue>
</Reference>
</SignedInfo>
<SignatureValue>
...
</SignatureValue>
<KeyInfo>
<X509Data>
<X509Certificate>
...
</X509Certificate>
</X509Data>
</KeyInfo>
</Signature>
</QIFDocument>
```

6.2.2 Use Case

Our goal is to demonstrate the benefits of digitally signing product data to support and improve data flows and quality control in a smart-manufacturing environment. Our use case focuses on manufacturing data elements transformations. Smart manufacturing requires digital product information to be available to each of its processes. The model-based paradigm on which smart manufacturing relies often needs product information to be expressed in different formats and processed by different software through the product lifecycle. This results in manipulating and generating variations of master-product models. These variations come from different transformations, either from the master models themselves or from other variations. These variations constitute what is identified as a transformation network, as seen in Figure 6.1.

A transformation network can be represented as a directed graph where nodes represent data and directed edges represent transformations. The head of a directed edge is the result of the transformation; the tail is the source to which the transformation is applied. Because of the significant number of transformations and variations during the product lifecycle, it is crucial to know who did what to whom and when. Our objective is to leverage X.509 certificates and digital signatures to embed and secure traceability information into the product models. This traceability information is crucial in a space where data is created and used on different platforms and in different locations by different users. Embedding and signing such information increases product data trustworthiness and enables reliable PDQ control. PDQ controls can take different forms. Quality control ensures data consistency, helps in troubleshooting data loss, identifies defective data processing, and improves the overall quality of data flows through the product lifecycle.

Four types of relevant traceability information are identified : i) who: identifies the system or person responsible for the generation of the attached dataset; ii) what: describes the type of

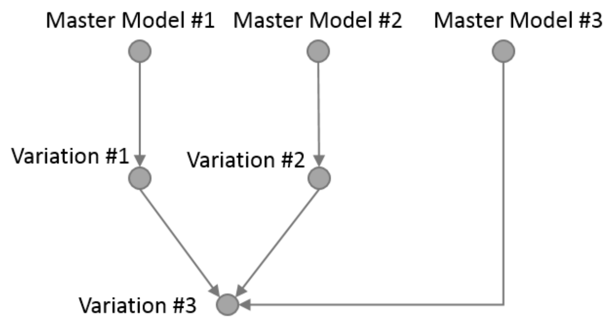


Figure 6.1: Transformation network

transformation that has been applied to a source to generate the attached dataset; iii) whom: describes the source data to which the transformation was applied to generate the attached dataset; and iv) when: time at which the transformation was recorded. Put back in the context of the transformation network, for any node at the head of a directed edge, embed the edge and its tail, the creation time, and the creator of the edge. Traceability information quality is vital to enable reliable PDQ control and requires consistency and accuracy. While the first, third, and fourth (who, whom, when) data fields are unambiguous, the second one (what) can be a source of ambiguity and generate inaccurate and ambiguous traceability information. To mitigate the possible ambiguity, the three most common transformations that occur during the product lifecycle are identified : i) translation: happens when the same information is reproduced in a different representation/file format; ii) inference: happens when computation or reasoning is applied on existing data to validate it or infer a new one; iii) augmentation: happens when an inference is run over a set of data and the result is added to the original set of data.

To illustrate the notion of a transformation network a trivial example (Figure 6.2) commonly found in a smart manufacturing environment was built. A native 3D computer-aided design (CAD) model was translated into a signed STEP AP242 model. That model was translated and augmented into a signed QIF MBD model. This new model, together with a list of metrology resources (QIF resources) and metrology knowledge (QIF rules), is reasoned on to generate a list of QIF measurement plans.

6.2.3 Extending ISO 10303-21 to support transformation network and multi-path hierarchical signings

Despite its current extension to support digital signatures, STEP Part 21 (10303-21 edition 3) cannot embed and sign traceability information such as the ones mentioned in Section 6.2.2 . Described is how ISO 10303-21 edition 3 embeds digital signatures into STEP Part 21 files. The main goal is to enrich digital signatures with meta-data about the provenance of the file content. This also improves the multiple-signatures support in STEP. In Part21

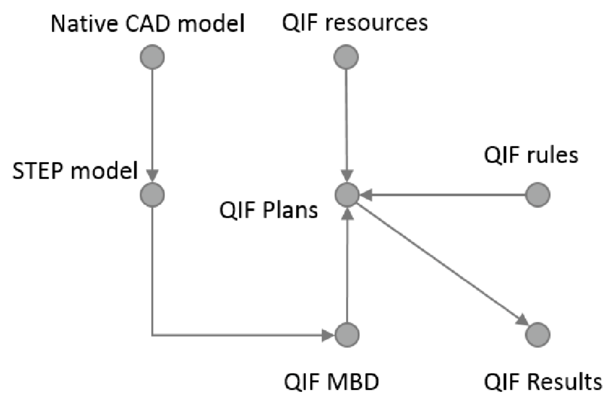


Figure 6.2: Example of a transformation network

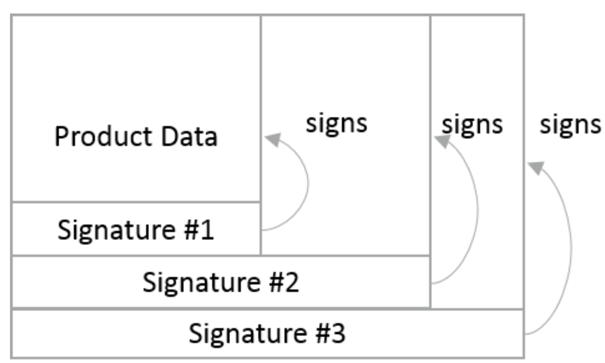


Figure 6.3: Multiple signatures support in STEP 10303-21 edition 3

edition 3, each signature section contains a digital signature that vouches for all the bytes before the section, whether those bytes are product data by itself or product data and other signatures. As shown in Figure 6.3, the signature mechanism requires Signature #2 to vouch for Signature #1, and Signature #3 to vouch for #2 and #1 combined. This mandatory vouching can raise legal issues and constraints in a field where product models and data must be legally signed and endorsed by a multitude of different actors from different organizations, along the product lifecycle.

ISO 10303-21 edition 3, in its current form, does not adequately support the business need for digital signature in an environment where it is a legal matter for companies to make sure their signatures are properly used and endorse the right content. Our goal is to enhance ISO 10303-21 to support multi-path hierarchical signing as shown in Figure 6.4. In a multi-path hierarchical signing, a new signature does not necessarily have to vouch for the latest signature in the file. This allows multiple organizations to sign on a same file while only vouching for signatures issued from their own organizations.

These enhancements to ISO 10303-21 require us to extend its syntax to support the following

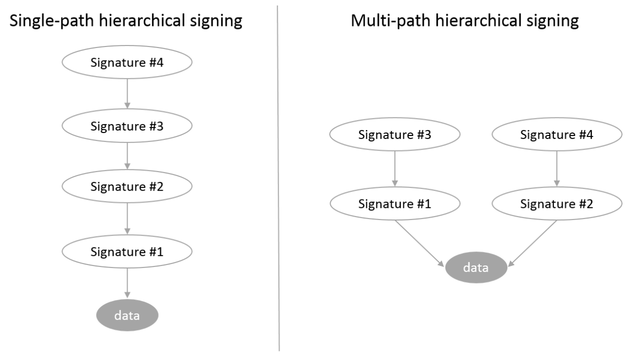


Figure 6.4: Single-path vs multi-path hierarchical signing

two requirements:

- A signature block must enable a signer to vouch for other existing signatures to support multi-path hierarchical signings
- A signature block must enable a signer to attach a set of metadata to document the signature and represent the file transformation information

ISO 10303-21 uses a Wirth Syntax Notation (WSN) [20] meta-syntax to define its syntax. A meta-syntax, or grammar, is a set of rules that describe and constrain a domain specific language and its valid syntax and vocabulary. Our WSN extension to the current WSN grammar is shown in Listing 6.4.

Listing 6.4: WSN extension to current WSN grammar

```
(1.1) SIGNATURE_SECTION = ``TRACE:''ENTITY_INSTANCE_NAME
TRACE
SIGNATURE
``ENDSEC;''.
(1.2) TRACE = ENTITY_INSTANCE_NAME ``=PKCS_TRACE(`` METADATA'')''.
(1.3) METADATA = ``{`` LIST_OF_FIELDS ''}''.
(1.4) LIST_OF_FIELDS = FIELD { ``,'' FIELD}.
(1.5) FIELD = FIELD_NAME``:''FIELD_VALUE.
(1.6) FIELD_VALUE = ``''' STRING ``'''.
(1.7) FIELD_NAME = STRING.
(1.8) SIGNATURE = ENTITY_INSTANCE_NAME ''=``PKCS_TOKEN .
(1.9) PKCS_TOKEN = ``PKCS(`` PKCS_SIGNATURE '',`` CROSS_BOOL '',``
    CROSS_INDEX'')''.
(1.10) PKCS_SIGNATURE= `` ' ''PKCS7`` ' ''.
(1.11) PKCS7 = STRING.
(1.12) CROSS_INDEX = ``[``LIST_OF_TRACE_IDS'']''
(1.13) LIST_OF_TRACE_IDS = ENTITY_INSTANCE_NAME {``,'' ENTITY_INSTANCE_NAME}.
(1.14) CROSS_BOOL = ``Y''|``N''.
```

Every signature block (Listing 6.4, 1.1) starts with the keyword `TRACE` followed by a numerical identifier in the form of an `ENTITY_INSTANCE_NAME`, ends with the delimiter `ENDSEC;` and contains two elements, a `TRACE` (Listing 6.4, 1.2) and a `SIGNATURE` (Listing 6.4, 1.8). A `TRACE` records `METADATA` (Listing 6.4, 1.2 and 1.3), as a set of `FIELD`s (Listing 6.4, 1.5) representing transformation information in a simplified JavaScript Object Notation (JSON) like format. The `SIGNATURE` (Listing 6.4, 1.8) itself contains a `PKCS7` (Listing 6.4, 1.10 and 1.11) string that represents its `PKCS7` encoding. This string is followed by a field – `CROSS_BOOL` (Listing 6.4, 1.9 and 1.14) – that determines whether the signature vouches for others or not. The last field – `CROSS_INDEX` (Listing 6.4, 1.9 and Listing 6.4, 1.12) – records the list of signature block identifiers that the current signature block vouches for.

The code snippet in Listing 6.5 is an example of our extension in use. In this example, a STEP file is signed three times (Listing 6.5, 2.3, 2.8, and 2.12) and contains three signature blocks (Listing 6.5, 2.1, 2.6, and 2.10). The first signature asserts that the current file results from a translation from December 2015 (Listing 6.5, 2.2). The second signature (Listing 6.5, 2.8) on the file certifies the authenticity of the STEP file itself and validates (Listing 6.5, 2.7) the transformation it comes from (Listing 6.5, 2.2). The third signature (Listing 6.5, 2.12) signs the STEP file. The associated metadata (Listing 6.5, 2.11) is used to record that the signer only acknowledges that the product and manufacturing information (PMI) in the STEP file conforms to the PMI in the native file.

Listing 6.5: Signature and Traceability example in a STEP document

```
(2.1) TRACE:#122
(2.2) #123 = PKCS_TRACE({source:`c:\\file.native', date:`12-DEC-2015',
    operation:`translation'})
(2.3) #124 = PKCS(`pkcs7_signature',N,[])
(2.4) ENDSEC;
(2.6) TRACE:#125
(2.7) #126 = PKCS_TRACE({source:`c:\\file.native', date:`12-DEC-2015',
    operation:`validation'})
(2.8) #127 = PKCS(`pkcs7_signature',Y ,[#122])
(2.9) ENDSEC;
(2.10) TRACE:#128
(2.11) #129 = PKCS_TRACE({source:`c:\\file.native', date:`12-DEC-2015',
    operation:`validation', what:`PMI'})
(2.12) #130 = PKCS(`pkcs7_signature',N,[])
(2.13) ENDSEC;
```

6.2.4 Extending QIF to support transformation network and multi-path hierarchical signings

QIF 2.1 XML implementation presents at the moment some of the same limitations as ISO 10303-21 edition 3. QIF does support multiple signatures, but its implementation is different from what Part 21 uses. It does not allow to attach data to signatures, which makes

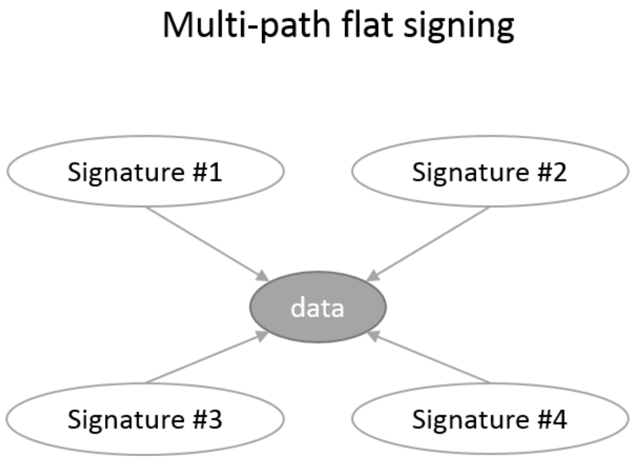


Figure 6.5: Mutli-path flat signing

it impossible to record the metadata that represents the transformation information.

The current version of the standard implements multiple signatures in a way that each signature vouches for the XML document alone, not the existing signatures at the time of signing, unlike Part 21. This pattern is categorized as a multi-path flat signing strategy, shown in Figure 6.5, because of the lack of hierarchy between the signatures. This section discusses an extension to the QIF information model to support a multi-path hierarchical signing strategy as shown in Figure 6.4.

QIF architecture relies on a `QIFDocument` container element that contains other elements, each representing concrete information (statistics, measurement plans, …). Section 6.2 discussed the `Signature` element owned by a `QIFDocument`. A summary of this architecture is shown in the UML class diagram in Figure 6.6. Not only does this architecture lack a placeholder for transformation and traceability metadata, it also requires any signature to vouch for all of the information contained in `QIFDocument`. This architecture was extended to support the same two requirements previously defined for Part 21: i) implementation of a multi-path hierarchical signing strategy and ii) a means to attach metadata to a digital signature. A new type was created – `Trace` – to represent a signature block. This signature block has a unique identifier `ID`, and contains two elements. The first element, of type `PKCS_TRACE`, is defined in a simplified JSON-like format that contains a unique identifier – `PKCS_TRACE_ID` – and a string – `Metadata` – used to record the traceability information, attached to the signature in the current signature block. The second element – `PKCS`, with a unique identifier – `PKCS_ID`, contains the digital signature itself. The `Cross_bool` attribute indicates whether the current signature and signature vouches for other signature(s) or not. If so, the `Cross_index` attribute lists the `QPIDs` of the signature block(s) vouched for by the current block. A summary of the extension is presented in the UML class diagram in Figure 6.7.

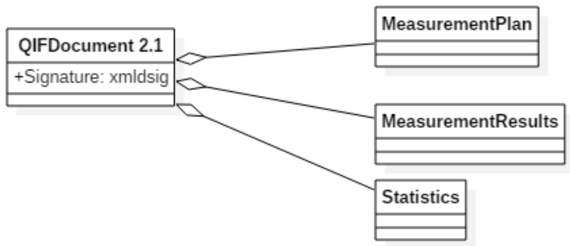


Figure 6.6: Digital Signature implementing in QIF 2.1

Listing 6.6 below shows an example of a QIF XML document that contains statistical information. The statistical information element has been digitally signed, independently from the rest of the document.

Listing 6.6: Signature and Traceability example in a QIF document

```
<?xml version="1.0" encoding="UTF-8"?>

<QIFDocument>
<QIFStatistics>
...
<Trace ID="TID1NIST">
<PKCS_TRACE PKCS_TRACE_ID="PTI1NIST">
<Metadata></Metadata>
</PKCS_TRACE>
<PKCS PKCS_ID="PID1NIST" Cross_bool="false">
<Signature>
...
</Signature>
</PKCS>
</Trace>
</QIFStatistics>
...
</QIFDocument>
```

6.2.5 Example usage of authentication, authorization, and traceability information

Recall, the approach is adding authentication, authorization, and traceability information as meta-data attached to the digital signature. This allows us to make a declaration of the quality of the data in a workflow similar to the Kikuchi et al. [21] usage scenarios discussed in Section 3.2.3. How the data may be used could also be declared. For example, assume there is a pre-defined verification criteria for data that would be used in a development or prototyping workflow. The information could be embedded in digital certificates to show that product data meets the PDQ requirements for development workflows and that the product data can only be used in a development workflow.

Figure 6.8 shows the example authentication, authorization, and traceability certification process for a development aerospace part. The first step in the example is to declare that the product data is of a development type. Next, an independently developed development-usage-PDQ-verification criteria is selected for checking that the product data is of a sufficient level to meet a development workflow's requirements. In this example, the aerospace product data is checked against PDQ criteria using a commercially available verification and validation system. The results of the verification (e.g., pass, warning, error) are captured and combined with the digital signature of the user running the verification check. The signature could also be from a system that is running the check autonomously. The resulting digital certificate containing the verification results, digital signature, and meta-data for authentication, authorization, and traceability is embedded in the aerospace part's product data. Listing 6.7 shows the authentication, authorization, and traceability information embedded in a STEP file translated and validated from the verified native product data.

Listing 6.7: Authentication, Authorization, and Traceability information in a STEP document for an aerospace part

```
TRACE:#3415
#3416 = PKCS_TRACE({source:`URI:15.1115\734.13.wingrib', date:`17-FEB-2016',
    operation:`verification', usage:`development', result:`pass with warnings',
    report:`URI:15.1115\734.13.wingrib.verification'})
#3417 = PKCS(`pkcs7_signature',N,[])
ENDSEC;
TRACE:#3418
#3419 = PKCS_TRACE({source:`URI:15.1115\734.13.wingrib',
    destination:`C:\\wingrib.stp', date:`17-FEB-2016', operation:`translation'})
#3420 = PKCS(`pkcs7_signature',Y,[#3415])
ENDSEC;
TRACE:#3421
#3422 = PKCS_TRACE({source:`c:\\wingrib.stp', date:`17-FEB-2016',
    operation:`validation', result:`pass', report:`c:\\wingrip-validation.pdf'})
#3423 = PKCS(`pkcs7_signature',Y,[#3415,#3418])
ENDSEC;
```

6.3 Conclusion

In this chapter I discussed how to leverage X.509-PKI-based digital certificates to restore trust in product data across the product lifecycle. This chapter showed that digital signatures are a means of authenticating, authorizing, and tracing product data. Embedding authentication, authorization, and traceability information in the product data builds trust throughout the product lifecycle.

During the effort I analyzed, evaluated, and implemented different commonly used standardized product-data formats. The use case highlighted gaps in the current version of these standards, gaps that I addressed for two of the formats: STEP and QIF. I am working with the appropriate standards-developing organizations to resolve the gaps and enhance each standard to fully support authentication, authorization, and traceability of product data.

The future efforts will focus on identifying gaps in other common standard formats and integrating digital certification of product data with various enterprise workflows. At the same time I recommend developing the metadata schema to embed with the certificates. I plan to work to integrate the DMC toolkit into a commercially available product-data management (PDM) tool to study automated processing of authentication, authorization, and traceability information for some of the most common enterprise workflows (e.g., engineering release, change management, manufacturing planning). I expect combining the DMC toolkit with automated enterprise workflow will significantly increase industry's confidence in product-data throughout the product lifecycle – such that industry can quickly understand who did what to whom and when it was done.

Chapter Bibliography

[1] Dazhong Wu, David W. Rosen, Lihui Wang, and Dirk Schaefer. Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. *Computer-Aided Design*, 59(0):1–14, 2015. ISSN 0010-4485. doi: http://dx.doi.org/10.1016/j.cad.2014.07.006. URL http://www.sciencedirect.com/science/article/pii/S0010448514001560.

[2] Dazhong Wu, Matthew John Greer, David W. Rosen, and Dirk Schaefer. Cloud manufacturing: Strategic vision and state-of-the-art. *Journal of Manufacturing Systems*, 32 (4):564–579, 2013. ISSN 0278-6125. doi: http://dx.doi.org/10.1016/j.jmsy.2013.04.008. URL http://www.sciencedirect.com/science/article/pii/S0278612513000411.

[3] Xun Xu. From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, 28(1):75–86, 2012. ISSN 0736-5845. doi: http://dx.doi.org/10.1016/j.rcim.2011.07.002. URL http://www.sciencedirect.com/science/article/pii/S0736584511000949.

[4] Thomas D Hedberg Jr, Nathan W Hartman, Phil Rosche, and Kevin Fischer. Identified research directions for using manufacturing knowledge earlier in the product life cycle. *International Journal of Production Research*, 55(3):819–827, 2017. doi: 10.1080/00207543.2016.1213453.

[5] Allison Barnard Feeney, Simon P. Frechette, and Vijay Srinivasan. A portrait of an ISO STEP tolerancing standard as an enabler of smart manufacturing systems. *Journal of*

Computing and Information Science in Engineering, 15(2):021001–021001, 2015. ISSN 1530-9827. doi: 10.1115/1.4029050. URL http://dx.doi.org/10.1115/1.4029050.

[6] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014.

[7] Energetics Inc. Measurement science roadmap for prognostics and health management for smart manufacturing system. Report, National Institute of Standards and Technology, 2015.

[8] Moneer Helu and Thomas Hedberg Jr. Enabling smart manufacturing research and development using a product lifecycle test bed. *Procedia Manufacturing*, 1:86–97, 2015. ISSN 2351-9789. doi: http://dx.doi.org/10.1016/j.promfg.2015.09.066. URL http://www.sciencedirect.com/science/article/pii/S2351978915010665.

[9] R. Gao, L. Wang, R. Teti, D. Dornfeld, S. Kumara, M. Mori, and M. Helu. Cloud-enabled prognosis for manufacturing. *CIRP Annals - Manufacturing Technology*, 64 (2):749–772, 2015. ISSN 0007-8506. doi: http://dx.doi.org/10.1016/j.cirp.2015.05.011. URL http://www.sciencedirect.com/science/article/pii/S000785061500150X.

[10] Min Li, Shuming Gao, and Charlie C. Wang. Real-time collaborative design with heterogeneous CAD systems based on neutral modeling commands. *Journal of Computing and Information Science in Engineering*, 7(2):113–125, 2006. ISSN 1530-9827. doi: 10.1115/1.2720880. URL http://dx.doi.org/10.1115/1.2720880.

[11] Asa Trainer, Thomas Hedberg Jr, Allison Barnard Feeney, Kevin Fischer, and Phil Rosche. Gaps analysis of integrating product design, manufacturing, and quality data in the supply chain using model-based definition. In *2016 Manufacturing Science and Engineering Conference*. American Society of Mechanical Engineers, 2016.

[12] World Wide Web Consortium. Semantic web, 2006. URL https://www.w3.org/standards/semanticweb/.

[13] Telecommunication Standardization Sector of ITU. Information technology – open systems interconnection – the directory – part 8: Public-key and attribute certificate frameworks, 2014. URL http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=64854.

[14] National Institute of Standards and Technology. Security and privacy controls for federal information systems and organizations. SP 800-53, 2015. doi: 10.6028/NIST.SP.800-53r4. URL http://www.nist.gov/manuscript-publication-search.cfm?pub_id=917904.

[15] M.J. Cheng and J.E.L. Simmons. Traceability in manufacturing systems. *International Journal of Operations and Production Management*, 14(10):4–16, 1994. doi: doi:10.1108/01443579410067199. URL http://www.emeraldinsight.com/doi/abs/10.1108/01443579410067199.

[16] M. H. Jansen-Vullers, C. A. van Dorp, and A. J. M. Beulens. Managing traceability information in manufacture. *International Journal of Information Management*, 23(5):395–413, 2003. ISSN 0268-4012. doi: http://dx.doi.org/10.1016/S0268-4012(03)00066-5. URL http://www.sciencedirect.com/science/article/pii/S0268401203000665.

[17] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 21: Implementation methods: Clear text encoding of the exchange structure, 2002.

[18] International Standards Organization. Document management – portable document format – part 1: PDF 1.7, 2008.

[19] International Standards Organization. Document management – 3D use of product representation compact (PRC) format – part 1: PRC 10001, 2014.

[20] Niklaus Wirth. What can we do about the unnecessary diversity of notation for syntactic definitions? *Commun. ACM*, 20(11):822–823, 1977. ISSN 0001-0782. doi: 10.1145/359863.359883.

[21] Yoshihito Kikuchi, Hiroyuki Hiraoka, Akihiko Otaka, Fumiki Tanaka, Kazuya G. Kobayashi, and Atsuto Soma. PDQ (product data quality): Representation of data quality for product data and specifically for shape data. *Journal of Computing and Information Science in Engineering*, 10(2):021003–021003, 2010. ISSN 1530-9827. doi: 10.1115/1.3402615. URL http://dx.doi.org/10.1115/1.3402615.

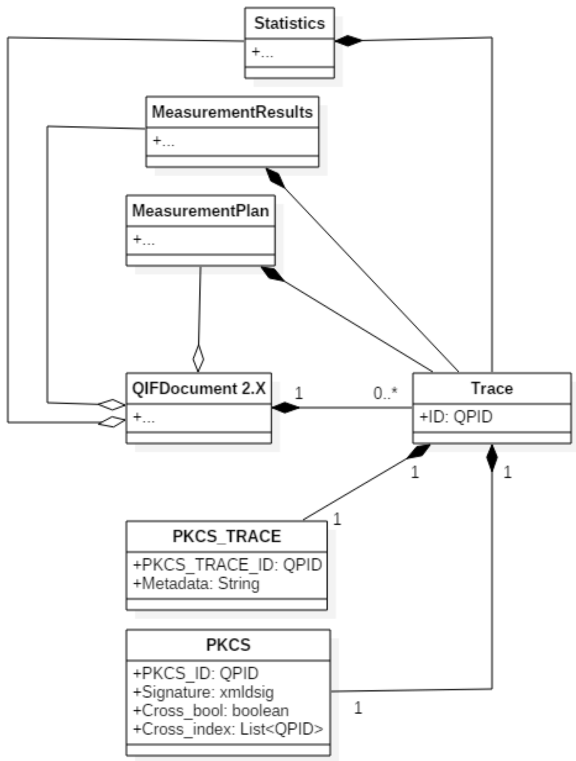


Figure 6.7: QIF extension for multi-path signing strategy support

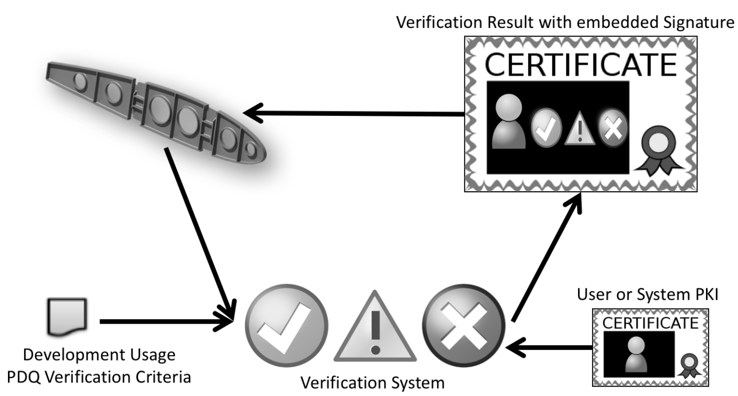


Figure 6.8: Example process for verifying the quality of product data and embedding usage restrictions

Chapter 7

A Product Lifecycle Built on Trust[1]

7.1 Introduction

Several software markets in the manufacturing sector, computer-aided design (CAD) in particular, are undergoing rapid evolution with the introduction of cloud, social, and mobile technology [1]. Evolving software technologies coupled with the trend [2, 3, 4] of digital and globally distributed manufacturing systems makes ensuring effective, efficient, and trusted product-data traceability an increasing challenge. Processes across the product lifecycle domains (e.g., engineering, manufacturing, quality assurance, sustainment) are often independent and disconnected – resulting in product-data traceability requiring a significant amount of human resources. Industry is shifting operations towards the concept of model-based enterprise (MBE) by changing the communication of product definitions and specifications from two-dimensional (2D) drawings (i.e., paper) to three-dimensional (3D) product representations (i.e., CAD models) [5].

The trends and challenges associated with the digital transformation of manufacturing increase the importance of product-data traceability and trustworthiness. Ensuring the necessary data is used by the correct function / role at the appropriate time is paramount, especially in regulated industries. This requires the ability to know that data being used is the correct type and version, is authorized for the intended usage, and came from the expected data owner / sender. In 2014, GrabCAD[2] surveyed [6] their current user base (at the time, over one-million users) to determine the most frustrating time-wasters in CAD collaboration. 75 percent of the survey respondents said they wasted time fabricating a prototype or production part using the wrong version of data. Other shocking statistics from the survey are: 80 percent of respondents spend time each month reconciling data versions because users were unaware of changes and 39 percent of respondents missed delivery dates waiting on data verification from their customer. It is the authors' opinion that the survey results point to a practical breakdown in the configuration management (CM) process.

Industry needs a dynamic way to trace and trust product data effectively and efficiently.

[1]This chapter was submitted as an article with the citation: Hedberg Jr, T., Krima, S., & Camelio, J. A. (In Review). Method for Enabling a Root of Trust in Support of Product-Data Certification and Traceability. *Journal of Computing and Information Science in Engineering*

[2]An online community of professional engineers, designers, manufacturers, and STEM students, `https://grabcad.com/`

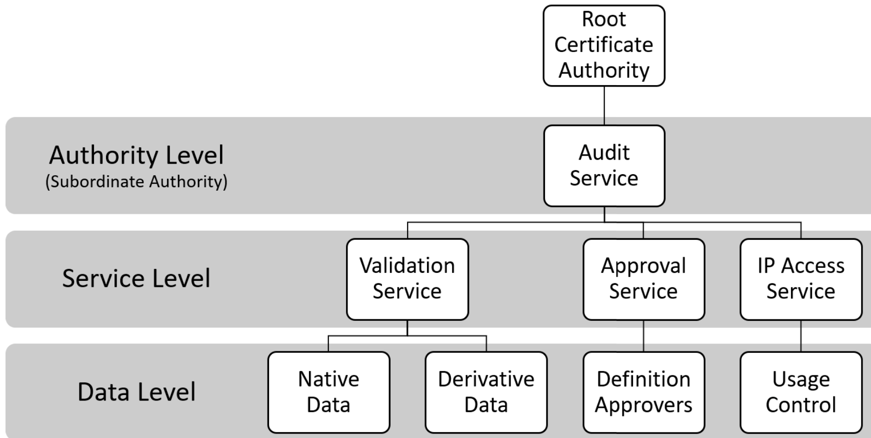


Figure 7.1: Hierarchy for chains of trust using X.509-PKI principles.

Previous research [7] shows how Public Key Infrastructure (X.509-PKI) from the X.509 standard [8] could be used to embed digital signatures into product data for the purposes of data certification and traceability. This paper will provide a review of technology that could be integrated to build trust throughout the product lifecycle. Then, the paper will propose a trust methodology supported by a structure and governance policy. Next, the paper will present a reference implementation and case study for common CM workflows typically found in regulated industries. Finally, the paper will draw conclusions and provide recommendations for further research to enable the product lifecycle of trust.

7.2 Trust Methodology

Trust is built on relationships. While federating trust forms more of a graph than a tree when visualized, trust still requires some structure and hierarchy to be effective and efficient. It is recommended that the three-tier hierarchy shown in Figure 7.1 to support explicit and implicit trust relationships. The top level of the trust structure is an authority level, the second is a service level, and the third is a data level. Using this type of structure would enable a trust environment based on authorization, authentication, and traceability as discussed in [7].

The authority level should not be considered the root certificate authority (CA). Instead, the authority level is an intermediate CA that has its certificates signed by a root CA and has the authority to sign certificates for entities the authority level chooses to trust. This approach aligns with X.509-PKI practices used today with digital signatures and Transport Layer Security (TLS) / Secure Sockets Layer (SSL) on the Internet. Visualized in Figure 7.2,

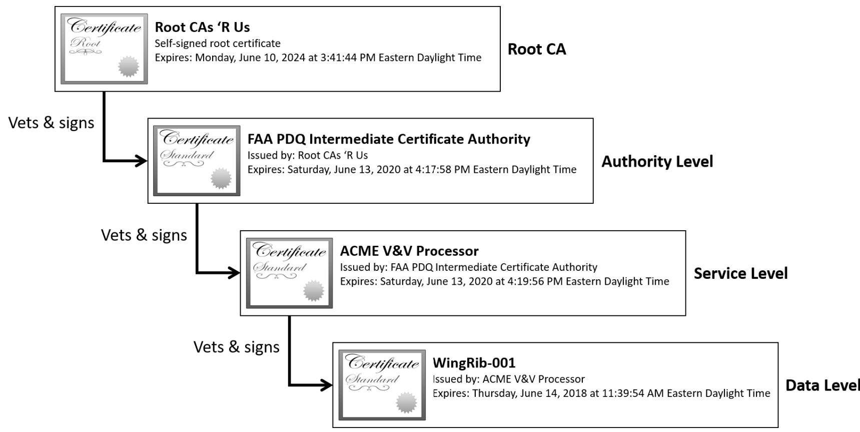


Figure 7.2: X.509-PKI certificates chain showing the levels of the trust structure.

a chain of trust begins with a root CA and links to the authority level, then to the service level, and finally to the data level.

Entities in the authority level act as an audit service – reviewing and certifying the service level entities. The ISO 9001 [9] audit process is a type of authority that could use digital certificates to create a chain of trust to provide evidence that an organization is ISO 9001 certified and to trace back to who certified the organization. Authority-level entities also include regulatory agencies (e.g., Federal Aviation Administration (FAA), Food and Drug Administration (FDA)), standards-accreditation organizations (e.g., International Standards Organization (ISO), American National Standards Institute (ANSI)), standards-development organizations (e.g., American Society of Mechanical Engineers (ASME), American Society for Testing and Materials (ASTM)), and manufacturers (e.g., Boeing Company certifying suppliers to their D6-51991 [10] quality specification). Regardless of the type of entity, the authority level should sign certificates of entities from only the service level.

Whereas ISO 9001 is a standard for quality management systems, ISO 16363 [11] is a standard that defines requirements for assessing the trustworthiness of digital repositories. The Primary Trustworthy Digital Repository Authorization Body Ltd. (PTAB) [12] is the first accredited organization to carry out audits in accordance with ISO 16363 and following procedures defined by ISO/IEC 17021-1:2015 [13], which defines requirements for bodies providing audit and certification of management systems. Authority level certificates could be used to build trust among organizations (e.g., PTAB) carrying out audits in accordance with ISO 16363 and ISO/IEC 17021. However, a long-term vision of any audit service should also include methods for auditing and certifying organizations, people, and products, too.

The authority level signs the service-level certificates, signaling that those service-level entities can be trusted if the authority is trusted. Three types of service-level actors are recommended: validation services, approval services, and intellectual property (IP) access services. The validation services audit or vet native and derivative data in the data level. The approval services audit or vet people and systems in the data level. Lastly, the access

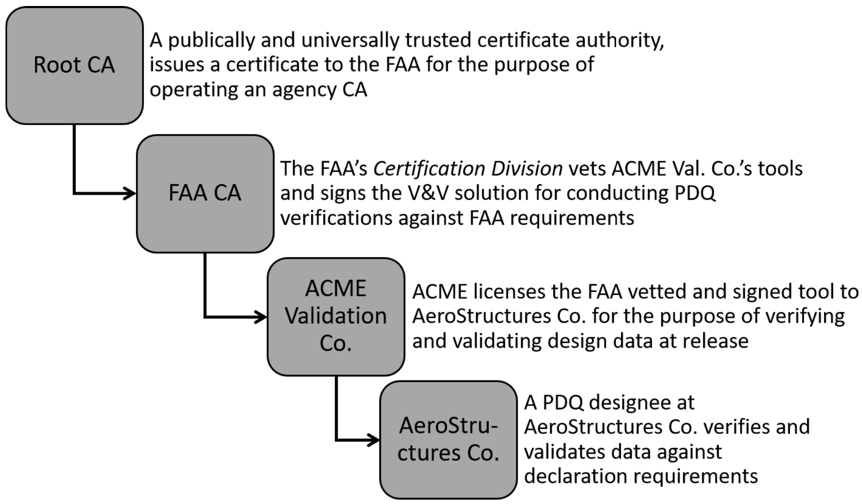


Figure 7.3: Chain of trust for a verification and validation system signing data with a digital signature

services set authorization permissions (e.g., who can use data, how data can be used) on the data. Section 7.2.1 through Section 7.2.3 provide more details and examples for interactions between the service level and data level.

7.2.1 Data Trust

Knowing where data came from and the quality of that data are key components to trustworthy communications and data exchange. There are several types of data (e.g., CAD models, analysis models, documents) that could be vetted by actors in the service level and signed with digital certificates. In the context of this chapter, the example is focused on the vetting of native CAD models (i.e., 3D models that come from an authoring CAD system) and derivative CAD models (e.g., 3D models that are translated from the native CAD system to a standard-based format such as Standard for the Exchange of Product Model Data (STEP) [14], Jupiter Tesselation (JT) [15], or Portable Document Format (PDF) [16] / Product Representation Compact (PRC) [17]).

Figure 7.3 shows an example chain of trust for a fictitious aerospace company using a fictitious verification and validation system to digitally sign data with digital certificates. In this scenario, a root CA signs a certificate issued to the FAA. The FAA vets a third-party verification and validation (V&V) solution from ACME Validation Company. ACME sells licenses of its V&V solution to customers who need to complete V&V workflows. AeroStructures Company licenses the software from ACME and uses V&V product-data quality (PDQ) criteria as declared requirements to check the quality of native and derivative CAD models produced by AeroStructures.

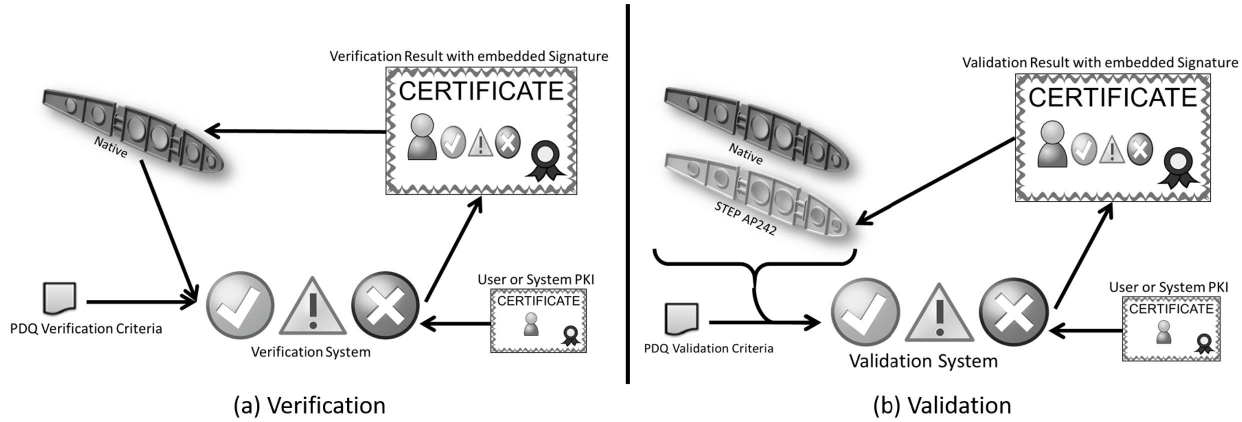


Figure 7.4: Example verification (a) and validation (b) workflows to ensure data meets a predefined level of quality (based on [7])

Further, consider a wing-rib produced by the AeroStructures Company. AeroStructures outsourced the fabrication of the wing rib. AeroStructures' supplier cannot use the native CAD model because the supplier uses a different CAD system than AeroStructures. Therefore, a derivative CAD model using the STEP standard is sent to the supplier. AeroStructures uses ACME's V&V software to verify and validate the CAD models before sending the STEP file to the supplier.

Figure 7.4 shows the verification and validation workflows that AeroStructures uses to complete the V&V tasks. In Figure 7.4(a), AeroStructures runs the ACME software on the native CAD model for the wing rib. Verification criteria that include data-quality rules and tolerances are used to configure the verification settings. After the quality of the model is confirmed, the results of the verification check are captured and AeroStructures is asked to sign the native model. The digital certificate attached to the native model includes the verification results and AeroStructures' digital signature. The same process is followed to complete the validation check on the derivative STEP data shown in Figure 7.4(b).

7.2.2 Person and System Trust

As shown in Figure 7.4, a person or system can digitally sign data. That signature could represent an approval of data or a confirmation that a task was completed. Examples of approvers are FAA designated-engineering representatives (DERs), company employees, and CM systems. Figure 7.5 presents a chain of trust for a person with a FAA DER certification who approves data in accordance to FAA requirements.

The FAA Aviation Safety Division appoints DERs. Under the trust structure, the FAA signs the certificate of John Doe, who is an employee of AeroStructures Company and has met the DER requirements. John Doe then uses his X.509-PKI-based certificate to digitally approve

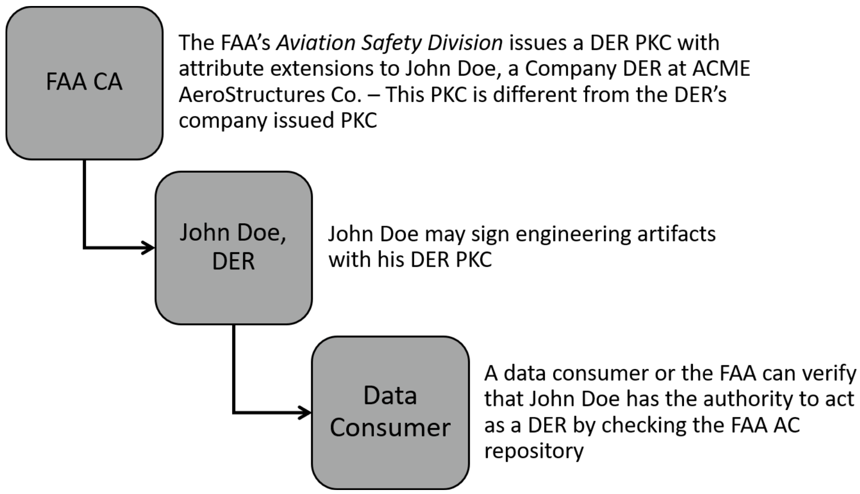


Figure 7.5: Verifying the chain of trust for a FAA DER signing data with a digital signature

data with a digital signature. John Doe's digital signature is verifiable to prove that John Doe has the authority to approve the data in accordance with FAA requirements. The full chain of trust is available to show that the FAA vetted John Doe and appointed him as a DER for the AeroStructures Company. Further, the chain of trust helps to determine that John Doe has the authority to approve the data that has his digital signature attached.

Thus, trust, in cases like the one shown in Figure 7.5, is represented semantically in data. Semantic representation of trust does not require interpretation of authorization like a hand-written signature. With digital signatures, the chain of trust either provides evidence of trust or it does not. The interpretation is binary.

7.2.3 Usage Trust

The concept of controlling the usage of digital files is not new. The use of digital rights management (DRM) became popular in the late-1990s and early-2000s as the music, movie, and publishing industries tried to curb pirating of their digital assets. Today, commercial solutions are becoming available in the manufacturing sector because of two main drivers. The first is the significant increasing focus on additive manufacturing. The second is IP protection of digital datasets as manufacturers increase usage of model-based definitions.

However, the current commercial applications are proprietary systems that require additional software and significant cost to deploy across the supply chain. In addition, the commercial solutions are not interoperable. Proprietary software must be installed for each commercial solution. This is problematic for the small-to-medium enterprise (SME) manufacturers who have to support multiple customers that may be using different DRM solutions.

The proposed approach described in this paper reduces the burden on the supply chain by

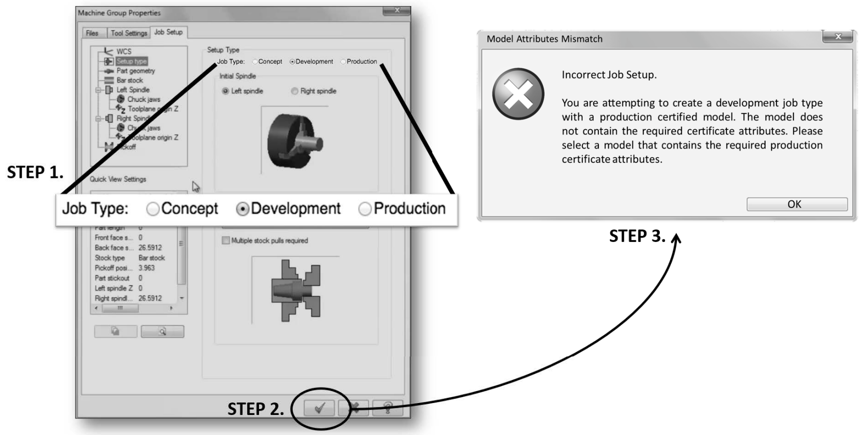


Figure 7.6: Example of data-usage rights controlling how data must be used in a CAM system

deploying an interoperable solution based on open-standards-based technologies. Data-usage authorization meta-data would be embedded in the X.509-PKI-based digital signature of a file. Should the authorization meta-data be changed or tampered with, the digital signature would become invalid. Figure 7.6 presents a use case showing how the data-usage rights could be controlled in a computer-aided manufacturing (CAM) system.

In the scenario depicted in Figure 7.6, the data owner decides to restrict a model to be used only for production purposes. When the manufacturing planner imports the model into the CAM system, he accidentally sets up the job as a development run (Figure 7.6, Step 1). When the manufacturing planner clicks the accept button to move to the next planning step (Figure 7.6, Step 2), the CAM system checks the authorization meta-data and determines that the model cannot be used as selected (Figure 7.6, Step 3). This is a basic example of managing the usage of data with X.509-PKI. However, the same approach could be deployed to other use cases, such as, but not limited to, controlling the number of times a part can be 3D printed, controlling who can access the file, and setting an expiration date for the model.

7.3 Tracing Transactions

As mentioned in Section 3.2.1, data traceability is paramount to enabling trustworthiness throughout the product lifecycle. Built on X.509-PKI and blockchain technologies, two use cases are defined for tracing data transactions throughout the lifecycle. A data transaction occurs anytime data ownership is declared or when data is exchanged between two actors. The first use case is file-only transactions, discussed in Section 7.3.1. The second use case is blockchain-registered transactions, discussed in Section 7.3.2.

Listing 7.1: Embedded *Verification*, *Release*, and *Revision* traceability information in a STEP document for an aerospace part

```
ISO-10303-21;
HEADER;
FILE_DESCRIPTION((`WingRib-001 for demonstration of trust and traceability
    in the product lifecycle'),`2;1');
FILE_NAME( `WingRib-001_rev01.stp',`2017-07-17T13:21:18',(`'),(`'),`',`',`');
FILE_SCHEMA((`AP242_MANAGED_MODEL_BASED_3D_ENGINEERING_MIM_LF { 1 0 10303
    442 1 1 4 }'));
ENDSEC;
DATA;
#1=APPLICATION_CONTEXT(`Managed model based 3d engineering');
.
.
.
TRACE:#3415
#3416 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    date:`2017-06-14T11:39:54', operation:`verification',
    usage:`production', result:`pass with warnings',
    report:`URI:15.1115\734.13.wingrib.verification'})
#3417 = PKCS(`pkcs7_signature_1',N,[])
ENDSEC;
TRACE:#3418
#3419 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    destination:`URI:15.1115\734.13.wingrib.release', usage:`production',
    date:`2017-06-20T15:18:36', operation:`release'})
#3420 = PKCS('pkcs7_signature_2',Y,[#3415])
ENDSEC;
TRACE:#3421
#3422 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib.release',
    destination:`URI:15.1115\734.13.wingrib.change01', usage:`production',
    date:`2017-07-17T13:21:18', operation:`revision'})
#3423 = PKCS(`pkcs7_signature_3',Y,[#3418])
ENDSEC;
END-ISO-10303-21;
```

7.3.1 File-Only Transactions

File-only transactions are asynchronous, require significant leveraging of X.509-PKI certificates, and require trust of the other actors with whom data is exchanged. Traceability is managed with metadata stored within the data files (See Listing 7.1). Figure 7.7 presents a use-case diagram for digitally signing a design specification and exchanging the file with another data user, using the digital manufacturing certificate (DMC) toolkit described in Chapter 6.

The `PKCS_TRACE` element on lines 13, 17, and 21 of Listing 7.1 contains meta-data that describes the context under which the data was digitally signed. At a minimum, the following attributes should be included in the `PKCS_TRACE`.

- `source:` identifies the source data that was reviewed and digitally signed, may be a circular reference to data containing the `PKCS_TRACE` element
- `destination:` identifies the actual signed data, which may be a circular reference to data containing the `PKCS_TRACE` element
- `usage:` the purpose(s) / use(s) for which the signed data is authorized
- `operation:` the reason why the data was signed (e.g., release, revision, validated)
- `date:` the date the operation was completed

Three actors are depicted in the file-only transactions use case: 1) data owner, 2) data user, and 3) bad actor. The data owner (herein owner) and data user (herein user) are the normal roles that would typically share data while executing tasks. When the owner is prepared to release the data to the user, the owner would review and sign the data using the DMC toolkit. Then, the owner would send the data to the user. The owner and user would store that signed data in their respective data repositories. The user would use the data to complete all agreed-upon tasks for the owner (e.g., supplier fabricates a part for a customer). This portion of the use case represents typical manufacturing-related business relationships.

Data could be compromised and/or stolen from owners and users by bad actors. In the file-only transactions use case, a bad actor could steal data from the user by compromising (e.g., gaining unauthorized access) to the user's data repository. The bad actor would have access to the signed data. If the owner somehow then found the signed data in the possession of an unauthorized actor, the owner could go back to his/her repository and determine all the users the data was sent to by querying and reviewing the certificate and metadata. This would provide the owner the ability to discover who received the data and request those users to investigate their systems for breaches. In this case, the owner would simply discover that he/she has a data problem, but the owner would not immediately know the root cause of that problem without further investigation.

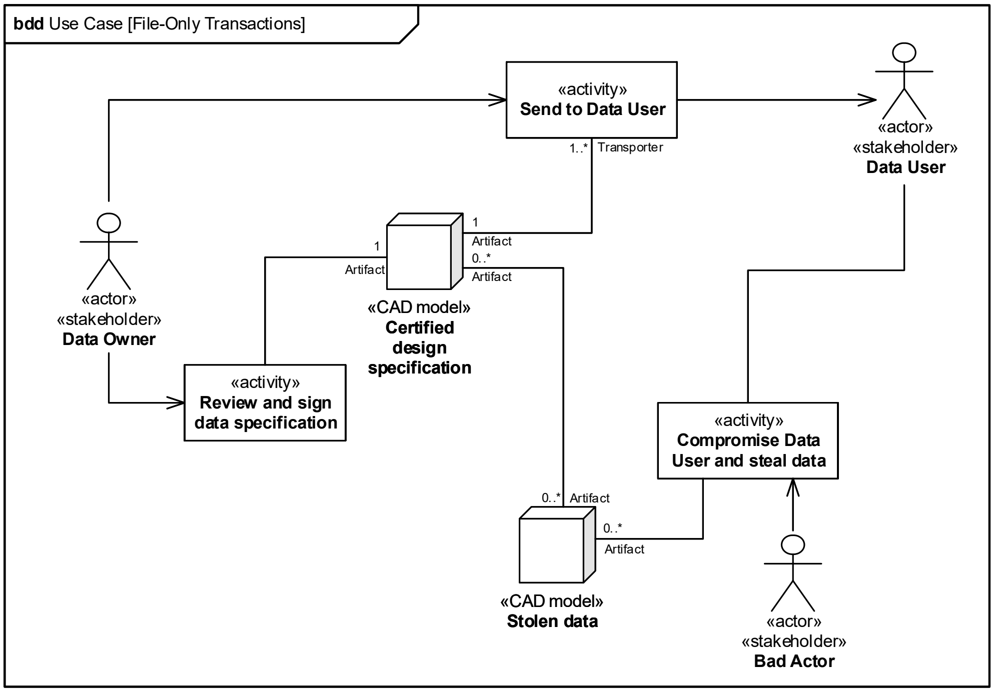


Figure 7.7: Use case for digitally signing a design specification (e.g., CAD model) and exchanging the data with another data user

However, the file-only transactions use case represents a solid foundation with which to build data-traceability principals and methods. Having the ability to quickly impart additional metadata into a file and then later be able to trace where the data came from, its purpose, and potential uses would reduce the risk of errors being introduced due to the wrong data being used or because of changes that went unnoticed.

7.3.2 Blockchain-Registered Transactions

Blockchain-registered transactions are synchronous and require leveraging X.509-PKI certificates and a blockchain (i.e., a distributed ledger). Traceability is managed with transactions registered in a blockchain. Figure 7.8 presents a use-case diagram for digitally signing a design specification, using the DMC toolkit described in Chapter 6, and registering data-ownership and data-exchange transactions in a blockchain.

The same three actors depicted in the file-only transactions use case are also depicted in the blockchain-registered transaction use case. The owner and user are still the normal roles that would typically share data between each other for the purposes of executing tasks. However, in this case, when the owner is prepared to release and send the data to the user, the owner would review and sign the data using the DMC toolkit and register the signature fingerprint in a blockchain to prove ownership of the data. Krima et al. [18] recommend storing only the

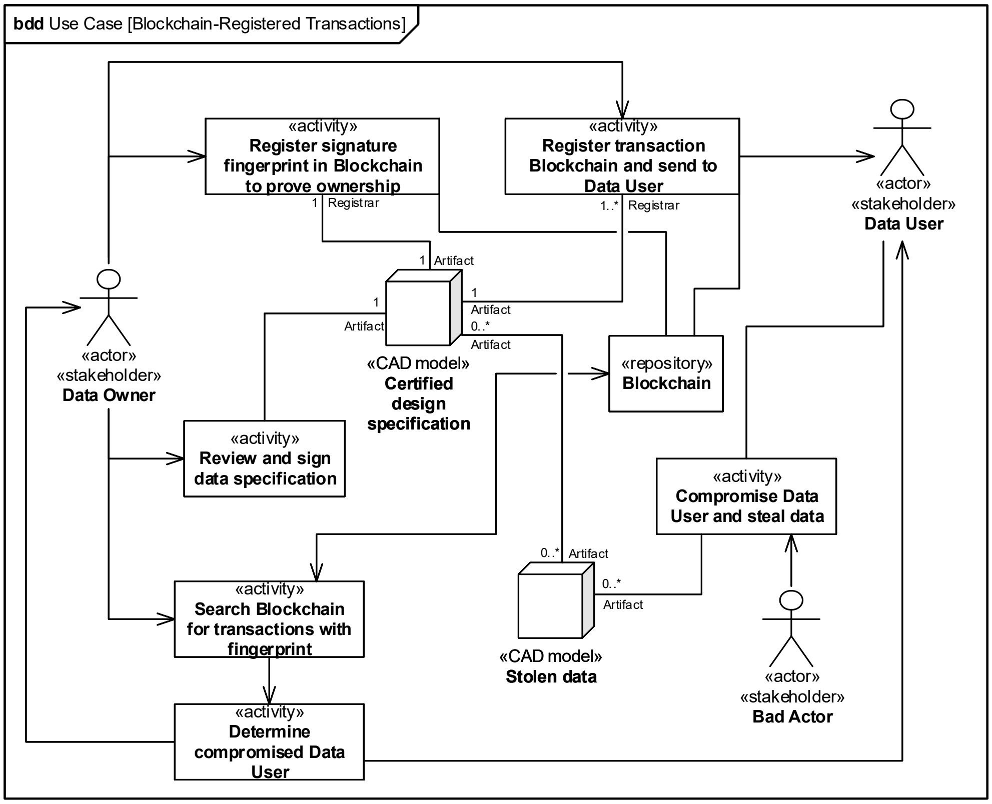


Figure 7.8: Use case for digitally signing a design specification (e.g., CAD model) and registering ownership and data-exchange transactions in a blockchain

signature fingerprint in the blockchain, registering the signature fingerprint in a transaction sent by the owner to him/herself for proving ownership, and then registering the signature fingerprint in transactions whenever the data is sent to a user. The owner and user would still store signed data in their respective data repositories. The user would also still use the data to complete all agreed-upon tasks for the owner (e.g., supplier fabricates a part for a customer). This portion of the use case, like the file-only transactions, represents typical manufacturing-related business relationships with the only difference being that each action on the data is registered in a blockchain.

The strength of the blockchain-registered transactions use case is in dealing with bad actors. If the owner found signed data in possession of an bad actor, the owner could query the blockchain and determine the exact transaction that was related to the compromised data. This provides the owner the ability to discover exactly who was authorized to receive the data originally and request that user to investigate his/her systems for breaches. In this case, the blockchain-registered transactions use case is differentiated from the file-only transactions use case because the owner would discover that he/she has a data problem and immediately know the root cause of the problem without further investigation.

7.4 Configuration Management Use Cases

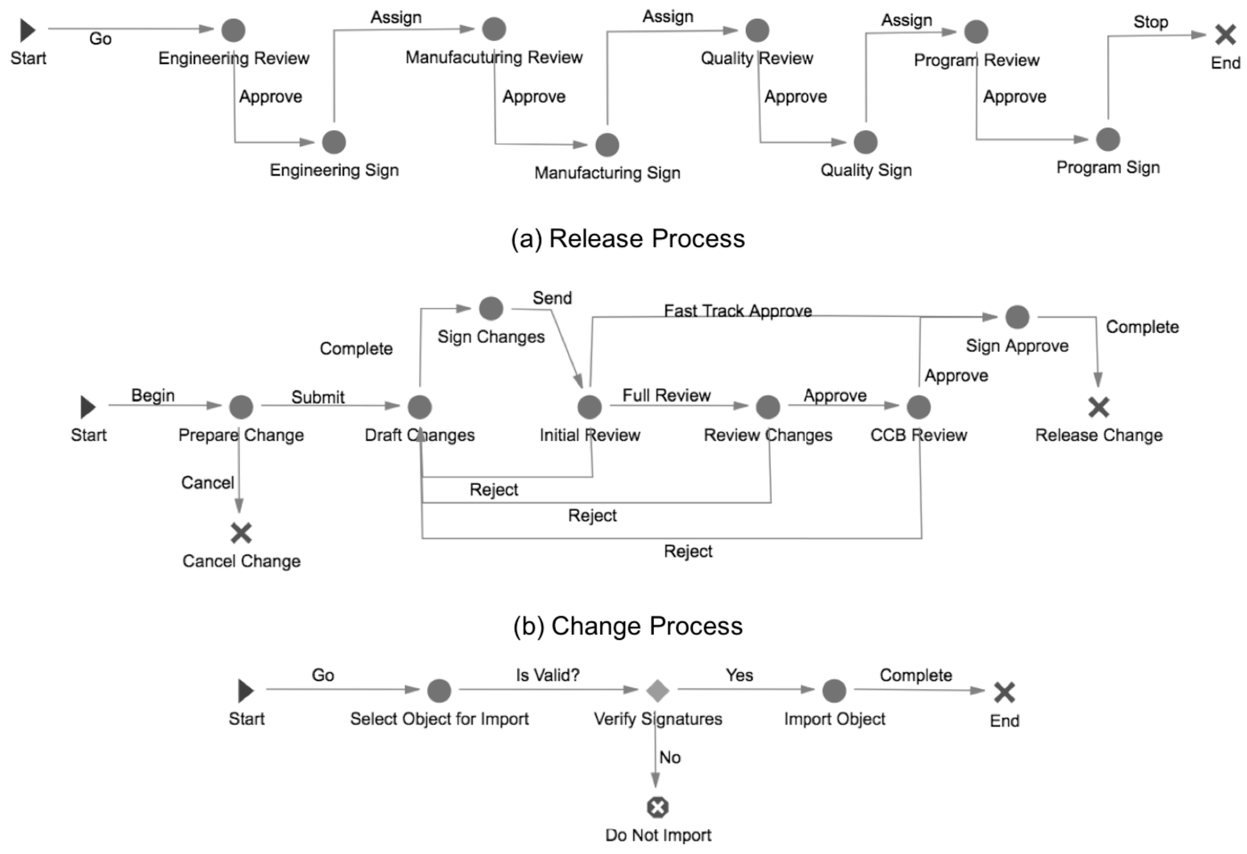
Ensuring the correct data is used when required by the needed functions and roles in the lifecycle depends on trusting the processes for making a product and managing the workflow. CM is key to managing the workflow and using the data. CM is a methodology for ensuring that a product's performance, function, and physical attributes are consistent with the requirements, design, and operational information of the product [19]. CM is especially crucial at the points data will released, changed, and/or exchanged. Each phase of the lifecycle must be cognizant of the CM of products because releases, changes, and process execution often happen in parallel as product specifications mature. Recall from the introduction that 75 percent of respondent manufacturers reported they have wasted time making a prototype or doing a production run using the wrong version of a CAD file [6]. Industry constantly struggles with using outdated versiona and product configurations. Ensuring the data configuration for a product specification is correct and being used appropriately is difficult today because manufacturing has become decentralized and distributed across the globe.

Using the method presented in this paper would address several CM issues facing industry. For example, the trust structure presented in Section 7.2 would assure that the right data came from an authorized person or system and that the data is used correctly. Metadata embedded in data would prove ownership, and certificate metadata and authorized digital signatures would allow determination of the validity of that data. Should software and tool providers adopt and implement this approach, industry would enjoy better control of data usage and could ensure that only data authorized for production is used in production processes. Figure 7.9 presents three CM workflows that would benefit from the usage of digital certificates.

7.4.1 Release Process

Figure 7.9(a) presents a typical release process workflow defined in a product-data management (PDM) or CM system. The purpose of the workflow is to manage the multi-stage review cycle of data that is ready to be authorized for use outside the owning organization. The data is considered "released" once the data has been "signed off" by the proper authority. In the context of this paper, the release process is defined as an engineering review, manufacturing review, quality review, and then a final program-management review. The data would be digitally signed by the functional role after each review and approval or rejection.

Listing 7.2 provides an example of the traceability metadata and digital signature that would be embedded in the data. The traceability data for the engineering review starts on line 12 of Listing 7.2, the manufacturing review on line 16, the quality review on line 20, and the program review on line 24. After the last approval, the CM system would sign the data. The traceability data for the CM system starts on line 28 of Listing 7.2. Also, notice the value of the `destination` attribute changed on line 29 after the CM system completed its



(a) Release Process

(b) Change Process

(c) Import and Verify Process

Figure 7.9: Product-data management workflows for signing data during a release cycle, change process, and import verification

task. This is because the status of the data changed in the system and was marked as "released." Finally, the CM system should cross-sign the data – essentially vouching for the digital signatures of the engineering, manufacturing, quality, and program reviewers. This is why `#3677,#3680,#3683,#3686` are included in square brackets at the end of the signature entry on line 30 of Listing 7.2.

Now that the data is marked "released," it is available to be exchanged with users who need the data to complete some tasks. Those users can use the digital certificates to verify the data is in fact released and is intended for production use. This workflow prototype shows how digital certificates could be implemented to ensure the correct data is used in processes across the lifecycle such as the example described in Section 7.2.3.

Listing 7.2: Embedded traceability information in a STEP document after going through configuration-management processes

```
ISO-10303-21;
HEADER;
FILE_DESCRIPTION((`WingRib-001 for demonstration of trust and traceability
    in the product lifecycle'),`2;1');
FILE_NAME( `WingRib-001_rev03.stp',`2017-07-20T16:21:18',(`'),(`'),`',`',`');
FILE_SCHEMA((`AP242_MANAGED_MODEL_BASED_3D_ENGINEERING_MIM_LF { 1 0 10303
    442 1 1 4 }'));
ENDSEC;
DATA;
#1=APPLICATION_CONTEXT(`Managed model based 3d engineering');
.
.
.
TRACE:#3677
#3678 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    destination:`URI:15.1115\734.13.wingrib', usage:`production',
    date:`2017-06-20T15:18:36', operation:`release approval'})
#3679 = PKCS('engineering_signature',N,[])
ENDSEC;
TRACE:#3680
#3681 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    destination:`URI:15.1115\734.13.wingrib', usage:`production',
    date:`2017-06-20T15:20:03', operation:`release approval'})
#3682 = PKCS('manufacturing_signature',N,[])
ENDSEC;
TRACE:#3683
#3684 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    destination:`URI:15.1115\734.13.wingrib', usage:`production',
    date:`2017-06-20T15:23:18', operation:`release approval'})
#3685 = PKCS('quality_signature',N,[])
ENDSEC;
TRACE:#3686
#3687 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    destination:`URI:15.1115\734.13.wingrib', usage:`production',
    date:`2017-06-20T15:35:34', operation:`release approval'})
#3688 = PKCS('program_signature',N,[])
ENDSEC;
TRACE:#3689
#3690 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib',
    destination:`URI:15.1115\734.13.wingrib.release', usage:`production',
    date:`2017-06-20T15:35:52', operation:`release'})
#3691 = PKCS('CM-system_signature',Y,[#3677,#3680,#3683,#3686])
ENDSEC;
TRACE:#3692
#3693 = PKCS_TRACE({source:`URI:20.500.11993\734.13.wingrib.release',
    destination:`URI:15.1115\734.13.wingrib.change01', usage:`production',
    date:`2017-07-20T16:21:18', operation:`change submission'})
#3694 = PKCS(`change_signature',Y,[#3689])
ENDSEC;
END-ISO-10303-21;
```

7.4.2 Change Process

Figure 7.9(b) presents a typical change process workflow defined in a PDM or CM system. The purpose of the workflow is to manage the multi-stage review cycle of proposed data changes. The change to the data is described by an "Engineering Change Request," which goes through a series of reviews for approval. In the context of this paper, the change request would be a package of data that includes the actual data being changed and supporting documentation for the change (e.g., problem reports, revision history).

The workflow shown in Figure 7.9(b) includes two opportunities for signatures. The first signature is applied by the change submitter and the second signature is applied by the change approver. In either signature, the actor could be a system or person. For each review, traceability metadata and a digital signature would be embedded in the data similar to the method described for the release process. Lines 32 to 35 of Listing 7.2 provide an example of the traceability metadata and the digital signature for the change-request submission. The traceability metadata informs the data user that the signature `operation` was for the purpose of changing the `source` data to the `destination` data. The change submitter should also cross-sign the data if release signatures exist to signify that the change has validated the previous version prior to recommending the change. The cross-signing is shown on line 34 of Listing 7.2. The change approver should also cross-sign the data and reference the change submission signature when the change is approved. The "Engineering Change Order" to initiate the requested change should only be issued after both the change submission and approval signatures are applied to the data.

7.4.3 Import and Verify File

Figure 7.9(c) presents a data "import and verify" process. Data should be verified before being storing the data in a repository. This is to ensure only verified data is stored. In the "import and verify" process, the user would select data to import into a system. The system would then check any digital signatures present in the data and verify the signatures are valid. If the signatures are valid, then the system would store the data in the repository. If the signatures are not valid, then the system would reject the data and not store it in the repository.

7.5 Discussion

In this chapter, data traceability using X.509-PKI is shown to be a viable option for supporting trustworthiness across the product lifecycle. Further, the proposed trust structure and hierarchy described in Section 7.2 establishes a governance policy for managing the traceability of data. Gol Mohammadi, et al. [20] conducted a survey of the literature to determine trustworthiness attributes to develop metrics for measuring trust. The authors determined 12 high-level attribute categories and further decomposed some of the categories with sub-categories. Table 7.1 analyzes and maps the method and governance policy described in this paper to the trustworthiness attributes presented by [20].

Overall, the work presented in this paper addresses the attributes for trustworthiness well. Five of the 32 metrics are considered out of scope, which leaves 27 metrics remaining for assessing this work. The method in the paper fully (black filled circle in the table) or partially (white filled circle in the table) addresses 97 percent (15 full, 11 partial) of the 27 metrics with only one metric not addressed.

Table 7.1: The analysis trustworthiness in product data using digital manufacturing certificates, proposed trust structure, and hierarchy

Main Attribute	**Sub-category**	**Addressed**	**Description of the method's congruence**
Security	Accountability	●	The use of X.509-PKI with embedded metadata ensures a person and/or system can be called upon to account for actions performed on data.
	Audit-ability & Traceability	●	The digital signatures generate a reliable and secure audit trail.
	Confidentiality		This sub-category is outside the scope of the data and is the responsibility of the data-managing systems.
	Integrity	●	The digital signatures would become invalid if the data changes or becomes corrupt. The X.509-PKI digital certificate have a high-level of integrity if generated in conformance to the standard. Therefore, both accidental and malicious alterations in the data are discoverable, which ensures the integrity of the data.
	Safety		This sub-category is outside the scope of the data and is the responsibility of the data-managing systems.
	Non-Repudiation		This sub-category is outside the scope of the data and is the responsibility of the data-managing systems.

LEGEND: ● = fully addressed. ○ = partially addressed. *Continued on next page*

Table 7.1 – *Continued from previous page*

Main Attribute	**Sub-category**	**Addressed**	**Description of the method's congruence**
Compatibility	Openness	●	The method takes advantage of the X.509-PKI standard, which is widely adopted in information technology and systems. Therefore, the method is open and transparent in how it works.
	Re-usability	●	The method is interoperable between information technologies and systems, while also allowing the metadata to be extensible to support multiple use cases and/or domains.
Configuration-related Quality	Change Cycle & Stability	●	The method meets an acceptable level of stability because it utilizes the widely-adopted X.509-PKI standard.
	Completeness		Future work is required to develop a complete standardized metadata schema to assist in defining semantic representation of the metadata such that it can be computer-processable.
Compliance		●	A governance policy that addresses most needs of regulated and non-regulated industries is proposed for using the method. The method also conforms to industry-led consensus-based standards.
Privacy		○	The method and governance policy partially address the ability to control the usage of private information. The method, being built upon X.509-PKI, supports controlling public and private keys. However, the method and governance policy must be used in concert with a properly configured data-management system to fully control the usage of private information.

LEGEND: ● = fully addressed. ○ = partially addressed. *Continued on next page*

Table 7.1 – *Continued from previous page*

Main Attribute	**Sub-category**	**Addressed**	**Description of the method's congruence**
Cost			Cost is considered out of scope for the work described in this paper. However, the use of standards should help in minimizing cost.
Data-Related Quality	Data Integrity	●	Human errors, malicious attacks, intentional data modifications, transmission errors, system/software bugs or viruses, and hardware malfunctions are discoverable because any change in the data would invalidate the digital certificates / signatures embedded in the data.
	Data Reliability	●	The provenance of the data is traceable. Further, usage (authorization) of the data is controlled by embedded metadata. Therefore, the correctness of the data used by the user is controllable.
	Data Timeliness		Timeliness is considered out of scope for the work described in this paper. Timeliness depends on the systems managing, transmitting, and monitoring the data.
	Data Validity	○	The method partially addresses data validity by enabling the capture of verification and validation results in the metadata. However, a trusted verification and/or validation system must be used to perform the analysis.
Dependability	Accuracy	●	The method is highly accurate if the certificates are created and managed in accordance with the X.509-PKI standard.
	Availability	●	The use of X.509-PKI are intended to operate in asynchronous environments. Therefore, the ability to verify and validate the certificates should be highly available.

LEGEND: ● = fully addressed. ○ = partially addressed. *Continued on next page*

Table 7.1 – *Continued from previous page*

Main Attribute	**Sub-category**	**Addressed**	**Description of the method's congruence**
	Failure Tolerance	•	X.509-PKI is a well-established and stable technology. There are mechanisms in the standard for handling failures
	Flexibility & Robustness	○	The method and governance policy partially address this sub-category. A standardized metadata schema is required to ensure the method is robust enough to handle changes in context. Further study into all industry sectors the method and governance policy may apply to is required. The method and governance policy are minimally viable for usage in highly regulated industries, such as aerospace, automotive, and medical devices.
	Reliability	○	Uncertainty exists as to how reliable the method can be without a standardized metadata schema that is semantically representable. However, the X.509-PKI technology has been proven to be sufficiently reliable.
	Scalability	○	The method and governance policy is intended to be extensible to enable scalability. However, further work is need in establishing a metadata schema that covers the minimum information required for all desired domains.
	Maintainability	•	Using well-established consensus-based standards ensures the highest level of maintainability as system undergo evolution.
Performance	Transaction Time	○	The method was designed to be as efficient as possible. However, the transaction time will also depend on several variables outside the control of the method.

LEGEND: • = fully addressed. ○ = partially addressed. *Continued on next page*

Table 7.1 – *Continued from previous page*

Main Attribute	**Sub-category**	**Addressed**	**Description of the method's congruence**
	Throughput	○	The method was designed to be as efficient as possible. However, the throughput will also depend on several variables outside the control of the method.
	Response Time	○	The method was designed to be as efficient as possible. However, the response time will also depend on several variables outside the control of the method.
Usability	Satisfaction	●	The method takes advantage of standards that are well-establish and widely adopted.
	Learn-ability		This sub-category is outside the scope of the data and is the responsibility of the data-managing systems.
	Effectiveness	○	The method and governance policy were studied in the context of a limited set of use cases. Therefore, further work is required to determine all domains where the method and governance policy would enable users to achieve the specified goals of this work.
	Efficiency of Use	○	The method was designed to be as efficient as possible. However, the efficiency of use will also depend on several variables outside the control of the method.
Correctness		○	The method conforms mostly to the utilized standards and specifications. The method is in full compliance with the X.509-PKI standard. However, extensions and/or modifications to several data-format standards are required to fully take advantage of the method and governance policy. [7] provide several modifications required.

LEGEND: ● = fully addressed. ○ = partially addressed. *Continued on next page*

Table 7.1 – *Continued from previous page*

Main Attribute	Sub-category	Addressed	Description of the method's congruence
Complexity		•	The complexity sub-category addresses inverse relationships such that more service fragmentation is typically considered less trustworthy than monolithic services. The method and governance policy presented in this paper were designed to take those inverse relationships into account. Using the X.509-PKI standard assists in managing the fragmentation by enabling the traceability of the data. Therefore, as data moves between systems and services, the digital certificates can be verified and validated to support trusting the data.

LEGEND: • = fully addressed. ○ = partially addressed.

7.6 Conclusion

The method presented in this paper shows promise for enabling a root of trust in support of product-data certification and traceability, thus supporting trustworthiness across the product lifecycle. The trust structure was designed to enable traceability of data; people and systems; and the usage and interactions between data, people, and systems. In addition, the presented method was studied using two types of use cases: 1) transaction tracing, and 2) common CM workflows. The method addresses all the use cases sufficiently. Lastly, the method was analyzed using a set of metrics for measuring trustworthiness. Overall, the method performed well in addressing most of the trustworthiness metrics, but additional required work was also identified.

Needed future work includes two areas of study. The first area is a complete metadata schema for defining all stakeholder-needed traceability information in the digital certificates. The second area is recommendations for extending the widely adopted data formats and standards to enable embedding digital signatures in a normalized manner. Addressing the future work would ensure the method fully satisfies the needs for data certification and traceability layer of the Lifecycle Information Framework and Technology (LIFT) concept as described in Chapter 4. Further, Helu et al. [21] proposes a reference architecture for integrating heterogeneous manufacturing systems, in which the authors recommend using STEP Application Protocol (AP) 242 [14], ISO 6983 (G code) [22], MTConnect [23], and Quality Information Framework (QIF) [24]. The method in this paper provides the ability to embed digital certificates in each of the domain-specific standards-based data discussed in [21]. Further, a complete metadata schema would also contribute to developing the Minimum Information Model [25, 26], which is the common and domain-specific information elements combined to represent the complete set of information required to effectively communicate to all functions and roles in the product lifecycle.

In closing, the contribution of this research is a novel method for using existing technologies to enable data certification and traceability in manufacturing. The method provides the infrastructure and guidance for transmitting the information (e.g., provenance, metadata) required to enable trustworthiness in the product lifecycle. The method would extend the use of X.509-PKI to enable trustworthy storage and exchange of data in manufacturing. The extension would support industry's needs in meeting traceability requirements from regulatory agencies such as FAA and FDA. A data user must know who did what to whom and when it was done. Therefore, product data must be guaranteed by an authority to a predefined level of data quality and trustworthiness if that information is to be used throughout the product lifecycle.

Chapter Bibliography

[1] Rakesh Sharma. The problems with reinventing CAD software. *Forbes*, 2013. URL `http://www.forbes.com/sites/rakeshsharma/2013/08/09/the-problems-with-reinventing-cad-software/`.

[2] Dazhong Wu, Matthew John Greer, David W. Rosen, and Dirk Schaefer. Cloud manufacturing: Strategic vision and state-of-the-art. *Journal of Manufacturing Systems*, 32 (4):564–579, 2013. ISSN 0278-6125. doi: 10.1016/j.jmsy.2013.04.008.

[3] Xun Xu. From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, 28(1):75–86, 2012. ISSN 0736-5845. doi: 10.1016/j.rcim.2011.07.002.

[4] Dazhong Wu, David W. Rosen, Lihui Wang, and Dirk Schaefer. Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. *Computer-Aided Design*, 59(0):1–14, 2015. ISSN 0010-4485. doi: 10.1016/j.cad.2014.07.006.

[5] Asa Trainer, Thomas Hedberg Jr, Allison Barnard Feeney, Kevin Fischer, and Phil Rosche. Gaps analysis of integrating product design, manufacturing, and quality data in the supply chain using model-based definition. In *2016 Manufacturing Science and Engineering Conference*. American Society of Mechanical Engineers, 2016.

[6] GrabCAD. Where did the time go? Report, GrabCAD, 2014. URL: https://resources.grabcad.com/time-go/.

[7] Thomas D. Hedberg Jr, Sylvere Krima, and Jaime A. Camelio. Embedding x.509 digital certificates in three-dimensional models for authentication, authorization, and traceability of product data. *Journal of Computing and Information Science in Engineering*, 17 (1):011008–011008–11, 2016. ISSN 1530-9827. doi: 10.1115/1.4034131.

[8] Telecommunication Standardization Sector of ITU. Information technology – open systems interconnection – the directory – part 8: Public-key and attribute certificate frameworks, 2014. URL http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=64854. ISO/IEC 9594-8:2014.

[9] International Standards Organization. Quality management systems, 2015. ISO/TC 176/SC 2, ISO 9001:2015.

[10] R. Dougherty. Quality assurance standard for digital product definition at Boeing suppliers. Report D6-51991, The Boeing Company, 2017, Accessed 2017-08-14. URL http://www.boeingsuppliers.com/dpd.html. Archived by WebCite at http://www.webcitation.org/6siPUecWG.

[11] International Standards Organization. Space data and information transfer systems – audit and certification of trustworthy digital repositories, 2012. ISO/TC 20/SC 13, ISO 16363:2012.

[12] PATB Ltd. Primary trustworthy digital repository authorisation body ltd, audit & certification, overview, 2017, Accessed 2017-09-01. URL: http://www.iso16363.org/iso-certification/overview/.

[13] International Standards Organization. Conformity assessment – requirements for bodies providing audit and certification of management systems – part 1: Requirements, 2015. ISO/CASCO, ISO/IEC 17021-1:2015.

[14] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014. ISO/TC 184/SC 4, ISO 10303-242.

[15] International Standards Organization. Industrial automation systems and integration – JT file format specification for 3D visualization, 2012. ISO/TC 184/SC 4, ISO 14306.

[16] International Standards Organization. Document management – portable document format – part 1: PDF 1.7, 2008. ISO/TC 171/SC 2, ISO 32000-1.

[17] International Standards Organization. Document management – 3D use of product representation compact (PRC) format – part 1: PRC 10001, 2014. ISO/TC 171/SC 2, ISO 14739-1.

[18] Sylvere Krima, Thomas Hedberg Jr, and Allison Barnard Feeney. Securing the digital threat for smart manufacturing: A reference model for blockchain-based product data traceability. Report AMS 300-6, National Institute of Standards and Technology, 2018.

[19] SAE International. Configuration management standard, 2011. URL https://www.sae.org/standards/content/eia649b/. EIA649.

[20] Nazila Gol Mohammadi, Sachar Paulus, Mohamed Bishr, Andreas Metzger, Holger Könnecke, Sandro Hartenstein, Thorsten Weyer, and Klaus Pohl. Trustworthiness attributes and metrics for engineering trusted internet-based software systems. In *Cloud Computing and Services Science*, pages 19–35. Springer International Publishing, 2014. ISBN 978-3-319-11561-0. doi: 10.1007/978-3-319-11561-0_2.

[21] Moneer Helu, Thomas Hedberg Jr, and Allison Barnard Feeney. Reference architecture to integrate heterogeneous manufacturing systems for the digital thread. *CIRP Journal of Manufacturing Science and Technology*, 19:191–195, 2017. ISSN 1755-5817. doi: 10.1016/j.cirpj.2017.04.002.

[22] International Standards Organization. Automation systems and integration – numerical control of machines – program format and definitions of address words – part 1: Data format for positioning, line motion and contouring control systems, 2009. ISO/TC 184/SC 1, ISO 6983-1:2009.

[23] MTConnect Institute. Mtconnect standard: Part 1 - overview and protocol, 2014. URL http://www.mtconnect.org/media/39542/mtc_part_1_overview_v1.3.pdf.

[24] Dimensional Metrology Standards Consortium. Part 1: Overview and fundamental principles in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014.

[25] Shawn P. Ruemler, Kyle E. Zimmerman, Nathan W. Hartman, Jr Thomas Hedberg, and Allison Barnard Feeney. Promoting model-based definition to establish a complete product definition. *Journal of Manufacturing Science and Engineering*, 139(5):051008–051008, 2016. ISSN 1087-1357. doi: 10.1115/1.4034625.

[26] Alexander M. Miller, Nathan W. Hartman, Thomas Hedberg Jr, Allison Barnard Feeney, and Jesse Zahner. Towards identifying the elements of a minimum information model for use in a model-based definition. In *International Manufacturing Science and Engineering Conference*, page V003T04A017, Volume 3: Manufacturing Equipment and Systems, 2017. American Society of Mechanical Engineers. doi: 10.1115/MSEC2017-2979.

Chapter 8

Conclusion

8.1 Summary

The goal of this dissertation was to develop an integration framework that brings all phases and systems of a product lifecycle together to enable efficient and effective measurements of the lifecycle in support of data-driven methods, specifically related to knowledge building, decision support, requirements management, and control.

The following research questions were proposed for answering as the first steps towards achieving the stated goals:

1. How can Graph Theory be used to dynamically generate inter-domain links of product data between the phases of the product lifecycle?

2. How can linked product data be managed to enable authentication and authorization as the data moves between domains?

3. What kind of management and/or governance policy could be established to ensure trustworthiness of linked product data?

A technology agnostic approach was pursued for dynamically generating links. Then, a demonstration was presented as a reference implementation using currently available technology. Requirements were gathering for trustworthiness related to interacting with product data. A data-traceability method was developed to. Then, a model was developed to propose a policy for enabling product-data trustworthiness. Lastly, the policy model was integrated with the dynamic links generation capabilities to ensure compliance. All methods were developed around open, consensus-based standards to increase the likelihood of scalability.

All three research questions were answered during the course of this research. This work provides a starting point for applying control methods at the enterprise level of the product lifecycle. Scalability of this work to the discrete manufacturing domain is evident. Further work is required to determine scalability to other manufacturing domains, such as batch and continuous processes. In conclusion, the answers to the three research questions discussed in this work provide demonstrated novel theoretical contributions to the product lifecycle management (PLM) domain.

8.2 Future Work

Uncertainty Quantification. Calculating uncertainty and variation in the product lifecycle was not addressed in this work. Uncertainty quantification is understood at the domain-specific levels, but how do those uncertainties aggregate up into a total uncertainty understanding for the entire product lifecycle? Also, what are the modes of variation in the product lifecycle and where are those models most likely to occur? To support applying control theory to the complete product lifecycle, one must be able to understand how uncertainty and variation across the product lifecycle relates. One must know how to predict the uncertainties of lifecycle phases, how those aggregate together, and how to identify where variation could be introduced. From there one must determine how to minimize the uncertainties and variation through decision-support systems and selection mechanisms.

Decisions in Project Management. The prototype implementations of the Lifecycle Information Framework and Technology (LIFT) show that proper management and technological innovation are critical for successful deployment of smart manufacturing. It is recognized that evaluation-based decision outcomes are subjective. Therefore, additional research in evaluation criteria and methods for implementing project management approaches and decision theories into LIFT is needed to ensure outcomes are objective. For example, how decision trees, heuristics, Markov chain, and Bayesian networks could assist in evaluating the outcome of decisions is of interest to industry. Also, additional research is needed to develop measures related to output-input relationships in decision making. The goal of these measures would be to provide an efficient way to determine the effectiveness and performance efficiency of each decision against an over-all goal. This would help determine if an overall goal needs to be revised in light of the work determined by the local goals or vice-versa.

Extending Traceability. Future work is needed for recommendations to extending the widely adopted data formats and standards to enable embedding digital signatures in data using a normalized method. That work would ensure the method fully satisfies the needs for data certification and traceability layer of LIFT as described in Chapter 4. Helu et al. [1] proposes a reference architecture for integrating heterogeneous manufacturing systems, in which the authors recommend using Standard for the Exchange of Product Model Data (STEP) Application Protocol (AP) 242 [2], ISO 6983 (G code) [3], MTConnect [4], and Quality Information Framework (QIF) [5]. The method in Chapter 6 provides the ability to embed digital certificates in each of the domain-specific standards-based data discussed in [1]. Further, a complete metadata schema would also contribute to developing the Minimum Information Model [6, 7], which is the common and domain-specific information elements combined to represent the complete set of information required to effectively communicate to all functions and roles in the product lifecycle.

Certificates Metadata. The future efforts need to focus on identifying gaps in other common standard formats and integrating digital certification of product data with various enterprise workflows. At the same time I recommend developing the metadata schema to embed with the certificates. Exemplar workflows and metadata were discussed in Chapter 7, but further definition is required to support industrial adoption. While one major computer-aided design (CAD) vendor has agreed to integrate the digital manufacturing certificate (DMC) method into their tools, more vendors need to integrate the DMC toolkit into a commercially available tools across the product lifecycle. I expect combining the DMC toolkit with engineering tools and automated enterprise workflow will significantly increase industry's confidence in product-data throughout the product lifecycle – such that industry can quickly understand who did what to whom and when it was done.

Dynamically Generating Connections. The Lifecycle Handler System (LHS) must provide capabilities to generate and register artifacts in the digital thread and to link them using connections. This includes, for example, generating design models from requirements (e.g., design synthesis), or generating simulation models and manufacturing process plans from design models, or registering new machines and machine configurations on a factory floor. Further, the LHS should enable automated generation of connections between artifacts when one is generated from the other. For instance, connections between design and manufacturing models are automatically generated when manufacturing models are generated from design models. This would overcome the manual creation of connections between artifacts that is laborious. Further research is required to enable the autonomous linking capabilities. Specifically, in near-real-time, how are all the links across enterprises tracked? Or how can inference systems be used to facilitate tracking of links?

Dynamic Linking and Querying. The LHS presented in Chapter 5 must provide capabilities to generate and register artifacts in the digital thread and to link them using connections. Future work is required in understanding how to monitor interactions and connections between data and store the links to enable product-lifecycle observation. Having dynamic linking reduces the burden on human resources for managing data and links. But also, common questions could be developed to assist in answering key inquiries that are important to measuring and understanding the product lifecycle. These common questions were called "Frequently Asked Queries" in Chapter 5. Industry would benefit from a reference library of graph-based queries that could be deployed to answer key questions across the product lifecycle. Combining a library of common queries with the methods described in Chapter 5 could significantly reduce the effort of human capital in making decisions by leveraging the capabilities of generating contextual graph-based viewpoints and quickly extracting actionable intelligence through knowledge generation.

8.3 Realized Impacts

A goal for any project should be to provide value the stakeholders. This would ensure the stakeholders can leverage the project outputs to be more competitive and improve quality of life. Impact can be defined here as making a substantial, positive external change directly enabled by project outcomes, resulting from adoption or use by external entities (e.g., industry, government agencies, society). Some impact was realized as the result of the work described in this dissertation.

Linked-Data. A small business software provider, based in Atlanta Georgia, adopted the linked-data method described in Chapter 5. The software provider implemented the methodology in the provider's latest commercial software release. The tool is used by large and medium enterprises in the Aerospace and Space sectors. Further, discussions are taking place with Open Applications Group (OAGi) and Open Services for Lifecycle Collaboration (OSLC) to investigate standardizing the linked-data methods.

Digital Certificates. Two standards development organizations (SDOs) adopted the DMC approach described in Chapter 6. ISO 10303-21:2016 [8] is the STEP standard that defines the EXPRESS language. The DMC approach for STEP described by the Wirth Syntax Notation (WSN) in Listing 6.4 is normalized in ISO 10303-21:2016. The QIF standard [5] also adopted and normalized the proposed extensions to the standard described in Section 6.2.4. Further, a large CAD vendor, based in Boston Massachusetts, added the DMC approach to their development roadmap for embedding digital certificates in the vendor's proprietary native file format. The vendor made this decision based on the urging of two large Aerospace companies.

Chapter Bibliography

[1] Moneer Helu, Thomas Hedberg Jr, and Allison Barnard Feeney. Reference architecture to integrate heterogeneous manufacturing systems for the digital thread. *CIRP Journal of Manufacturing Science and Technology*, 19:191–195, 2017. ISSN 1755-5817. doi: 10.1016/j.cirpj.2017.04.002.

[2] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 242: Application protocol: Managed model-based 3D engineering, 2014. ISO/TC 184/SC 4, ISO 10303-242.

[3] International Standards Organization. Automation systems and integration – numerical control of machines – program format and definitions of address words – part 1: Data for-

mat for positioning, line motion and contouring control systems, 2009. ISO/TC 184/SC 1, ISO 6983-1:2009.

[4] MTConnect Institute. Mtconnect standard: Part 1 - overview and protocol, 2014. URL http://www.mtconnect.org/media/39542/mtc_part_1_overview_v1.3.pdf.

[5] Dimensional Metrology Standards Consortium. Part 1: Overview and fundamental principles in quality information framework (QIF) – an integrated model for manufacturing quality information, 2014.

[6] Shawn P. Ruemler, Kyle E. Zimmerman, Nathan W. Hartman, Jr Thomas Hedberg, and Allison Barnard Feeney. Promoting model-based definition to establish a complete product definition. *Journal of Manufacturing Science and Engineering*, 139(5):051008–051008, 2016. ISSN 1087-1357. doi: 10.1115/1.4034625.

[7] Alexander M. Miller, Nathan W. Hartman, Thomas Hedberg Jr, Allison Barnard Feeney, and Jesse Zahner. Towards identifying the elements of a minimum information model for use in a model-based definition. In *International Manufacturing Science and Engineering Conference*, page V003T04A017, Volume 3: Manufacturing Equipment and Systems, 2017. American Society of Mechanical Engineers. doi: 10.1115/MSEC2017-2979.

[8] International Standards Organization. Industrial automation systems and integration – product data representation and exchange – part 21: Implementation methods: Clear text encoding of the exchange structure, 2016. ISO/TC 184/SC 4, ISO 10303-21.

Appendix A

Acronyms

2D two-dimensional 137

3D three-dimensional 1, 33, 35–37, 42, 44, 79, 121, 122, 126, 137, 140, 143, 173

ALM application lifecycle management 52, 53, 92

AMT Association for Manufacturing Technology 34

ANSI American National Standards Institute 34, 139

AP Application Protocol 33, 83, 157, 162

API application programming interface xv, 54, 77, 78, 102, 105, 122

ASME American Society of Mechanical Engineers 37, 100, 139

ASTM American Society for Testing and Materials 139

BDD block definition diagram xi, 99, 101

BOM bill of materials 70, 71, 97

C&R cause and remedy 82, 83

CA certificate authority 138–140

CAD computer-aided design xi, xii, 1, 29, 33, 35–37, 42–44, 70, 71, 74, 80, 83, 95, 97, 100, 102, 117, 126, 137, 140, 141, 146–148, 163, 164, 173

CAE computer-aided engineering 72

CAI computer-aided inspection 43

CAM computer-aided manufacturing xii, 33, 43, 70, 93, 95, 143

CAx computer-aided technologies 83

CFD computational fluid dynamics 72

CM configuration management 137, 138, 141, 148, 149, 151, 157

CMS coordinate-measurement system 33

CNRI Corporation for National Research Initiatives 75, 76

CPM Core Product Model 53, 54

CPM2 Revised Core Product Model 53–55

CPSC Consumer Product Safety Commission 37

cUAV configurable unmanned aerial vehicle xi, 99, 101, 114

DARPA Defense Advanced Research Projects Agency 29

DER designated-engineering representative xii, 141, 142

DFM design for manufacturing 74

DLA Defense Logistics Agency 28, 29

DMC digital manufacturing certificate 122–124, 133, 145, 146, 163, 164

DMSC Dimensional Metrology Standards Consortium 34

DNS Domain Name System 55

DoD U.S. Department of Defense 7, 28, 29, 73, 76

DOI Distributed Object Identifier 94, 108

DRM digital rights management 142

ECR engineering change request 79, 80, 82, 83

ERP enterprise resource planning 43, 52, 71, 77

FAA Federal Aviation Administration xii, 6, 8, 35–37, 139–142, 158

FAIR first article inspection reporting 100

FDA Food and Drug Administration 6, 8, 37, 139, 158

FEA finite-element analysis 72

GD&T geometric dimensions and tolerances 33

GID global identifier 55, 96–98, 102

HTML Hypertext Markup Language 30

HTTP Hypertext Transfer Protocol 20, 30, 34, 98

HTTPS Hypertext Transfer Protocol over Secure Sockets Layer 42

IIoT industrial internet of things 53

INCOSE International Council on Systems Engineering 99, 100, 114

IP intellectual property 42, 84, 94, 108, 139, 142

ISO International Standards Organization 40, 139

ISS International Space Station 28

IT information technology 2, 96

ITU-T Telecommunication Standardization Sector of the International Telecommunication Union 40

JSON JavaScript Object Notation xv, 102, 105, 113

JT Jupiter Tesselation 43, 70, 140

LHS Lifecycle Handler System xi, xiii, 94–99, 107, 108, 113, 163

LIFT Lifecycle Information Framework and Technology x, xi, 7, 8, 26, 69, 71–73, 75–80, 82–86, 95, 157, 162

LOI Lifecycle Object Identifier 78

MADE Manufacturing Automation and Design Engineering 29, 30

MBD model-based definition 33, 35, 70, 121, 126

MBE model-based enterprise 2, 20, 32, 33, 36, 80, 86, 121, 137

MBM model-based manufacturing 33

MBSE model-based systems engineering 2, 99, 114

MEDALS Military Engineering Data Asset Locator System 7, 29, 73, 76

MES manufacturing execution system 43, 52, 71, 77

MTC Manufacturing Technology Centre 99, 100, 114

NASA National Aeronautics and Space Administration 28

NC numerical control 33, 100

NIST National Institute of Standards and Technology 6, 53, 83–85, 99, 100, 114, 122

NNMI National Network of Manufacturing Innovation 84

NSF National Science Foundation 28

OAGi Open Applications Group 164

OASIS Organization for the Advancement of Structured Information Standards 40

ODI Open Data Institute 26, 27

OEM original equipment manufacturer 36

OOI Ocean Observatories Initiative 27, 28

OSLC Open Services for Lifecycle Collaboration 43, 164

OSTP Office of Science and Technology Policy 84

OT operational technology 2

OWL Ontology Web Language 31

PDF Portable Document Format 33, 70, 140

PDM product-data management 43, 52, 71, 77, 93, 98, 133, 148, 151

PDQ product-data quality 37–40, 73, 74, 122, 125, 126, 132, 140

PHM prognosis and health monitoring 85

PI principal investigator 28

PLCS Product Life Cycle Support 43

PLM product lifecycle management xiii, 4, 7, 8, 45, 52, 53, 69–71, 78, 92, 122, 161

PMI product and manufacturing information 33, 129

PRC Product Representation Compact 33, 70, 140

PSI Physical Science Informatics 27, 28

PTAB Primary Trustworthy Digital Repository Authorization Body Ltd. 139

QIF Quality Information Framework xii, xv, 7, 32, 34, 35, 71, 74, 79, 80, 82, 83, 93, 100, 123, 124, 126, 129–131, 133, 136, 157, 162, 164

QMS quality management system 43, 71, 77, 93

RDF Resource Description Framework 20, 30, 31

REST Representational State Transfer xv, 34, 98, 102, 105

RII receiving and incoming inspection xv, 100, 102, 104

S/MIME Secure/Multipurpose Internet Mail Extensions 40

SaaS software-as-a-service 42, 43

SAML Security Assertion Markup Language 40

SDO standards development organization 164

SFTP Secure File Transfer Protocol 42

SKOS Simple Knowledge Organization System 31

SLR systematic literature review 8, 14

SME small-to-medium enterprise 71, 77, 142

SMOPAC Smart Manufacturing Operations Planning and Control 85

SMS Test Bed Smart Manufacturing Systems Test Bed 6, 83, 84

SOA service-oriented architecture 40, 78

SPMM semantic-based product metamodel 54

SSL Secure Sockets Layer 40, 138

STEP Standard for the Exchange of Product Model Data 6, 33, 43, 70, 83, 140, 141, 157, 162, 164

STEP AP242 Standard for the Exchange of Product Model Data Application Protocol 242 32, 33

STL Stereolithography 70

SysML Systems Modeling Language xi, 43, 55, 99, 101

TDP technical data package 44

TLS Transport Layer Security 40, 138

TSM Total System Model 54, 55

UML Unified Modeling Language 43

URI Uniform Resource Identifier 20, 30, 97, 98, 113

V&V verification and validation 38, 39, 140, 141

W3C World Wide Web Consortium 30, 79

WSN Wirth Syntax Notation 128, 164

WWW World Wide Web 29–31

X.509-PKI Public Key Infrastructure xii, 36, 40, 42, 122, 123, 132, 138, 139, 141, 143, 145, 146, 151–158

X.509-PMI Privilege Management Infrastructure 40

XML Extensible Markup Language 29, 34, 100, 124, 129–131

XSD XML Schema Definitions 35

Appendix B

Glossary

Cyber-Physical System (CPS) Engineered systems that are built from, and depend upon, the seamless integration of computational algorithms and physical components.

Digital Thread An integrated information flow that connects all of the phases of the product lifecycle using an accepted authoritative data source (e.g., technical data package, three-dimensional (3D) CAD model)

Digital Twin An integrated model, enabled by the Digital Thread, that combined data from both the cyber-space and physical-space to mirror and predict "things" (e.g., activities, performance, outcomes, events) over the life of the model's corresponding physical twin.

Linked Data A method of publishing structured data so that it can be interlinked and become more useful through semantic queries.

Product Lifecycle A high-level activity model that starts with marketing, continues through design and manufacture, to selling, then support, and ends with the decommission and disposal or recycle of the product.

Product Lifecycle Management (PLM) The business activity of managing, in the most effective way, a company's products all the way across their lifecycles; from the very first idea for a product all the way through until it is retired and disposed of.

Stage Gate Process Used to describe a point in a project or plan at which development can be examined and any important changes or decisions relating to costs, resources, profits, etc. can be made.